Advances in Science, Technology & Engineering Systems Journal

ASTES

VOLUME 10-ISSUE 4|JUL-AUG 2025

www.astesj.com

ISSN: 2415-6698

EDITORIAL BOARD
Editor-in-Chief

Editorial

The expansion of digital transformation, advanced scientific methods, and innovative pedagogical strategies continues to reshape diverse fields of knowledge and practice. From education and artificial intelligence to atmospheric science, electrical engineering, and instructional reform, research efforts are converging to address pressing challenges while creating pathways for improvement and innovation. The following studies highlight significant contributions across these domains, illustrating how technology, models, and frameworks can enhance performance, accuracy, and adaptability in both academic and industrial contexts.

The quality of higher education in Brazil faces significant challenges arising from digital distractions, weak academic foundations, and high dropout risks. To address these issues, a machine learning-based assessment tool called AILA has been developed to predict academic performance and recommend personalized study resources. Using psychometric profiles of more than 41,000 students, the system provides tailored support in fundamental areas such as Portuguese and Mathematics. With CatBoost achieving an accuracy of 0.74 in predicting proficiency, the findings demonstrate how artificial intelligence can optimize individual learning trajectories, strengthen academic engagement, and support institutional strategies for improving student outcomes [1].

In regions with limited ground-based ionospheric monitoring infrastructure, satellite-based approaches are offering new opportunities for atmospheric science. Using radio occultation data from the COSMIC-2 mission, this study provides the first analysis of ionospheric peak parameters over equatorial Africa. Results show distinct diurnal and seasonal variations, with NmF2 values peaking after sunrise and reaching higher levels during equinoxes compared to solstices. A comparative analysis with the IRI-2016 model revealed significant overestimations, particularly during declining phases, with discrepancies reaching up to 60 percent. By highlighting these differences, the research underscores the value of satellite-based observations for improving ionospheric modeling in regions where data has been historically scarce [2].

In the study of electrical machines, the influence of MMF space harmonics on squirrel-cage induction motors has long been recognized, yet often fragmented in analysis. This work integrates the origin, behavior, and interaction of harmonics into a unified system that explains parasitic torques, magnetic force waves, and their physical relationships. A comprehensive equivalent circuit diagram and tabulated representation of magnetic forces provide clarity on synchronous and asynchronous phenomena, reducing the need to consider an infinite range of harmonics. By identifying the harmonics most relevant to design and performance, the study offers both theoretical and practical contributions, enabling more precise calculations and more efficient motor design methodologies [3].

Engineering education requires adaptive pedagogical models that align academic learning with the demands of rapidly changing industries. This research proposes a blended instructional framework based on micro-topic pedagogy under the New Engineering Education paradigm. In a quasi-experimental study involving 132 learners, the experimental group using the micro-topic approach demonstrated significant improvements in applied skills, participation, and problem-solving abilities compared to traditional methods. Real-time simulation tools and adaptive feedback mechanisms amplified these outcomes, with a 23 percent increase in troubleshooting proficiency. The results position this model as a scalable and effective strategy to strengthen interdisciplinary expertise and bridge the gap between academic training and professional application [4].

Together these studies underscore the transformative power of advanced technologies, systematic analysis, and pedagogical innovation across different fields. Whether enhancing learning outcomes in higher education, improving ionospheric modeling, optimizing electrical machine design, or reforming engineering education, these contributions provide practical solutions and theoretical advancements. They collectively highlight how interdisciplinary approaches and technological integration can drive progress in both scientific research and societal development.

References:

[1] G. Brás, S. Leal, B. Sousa, G. Paes, C. Junior, J. Souza, R. Assis, T. Marques, T.T.C. Silva, “Machine Learning Methods for University Student Performance Prediction in Basic Skills based on Psychometric Profile,” Advances in Science, Technology and Engineering Systems Journal, 10(4), 1–13, 2025, doi:10.25046/aj100401.

[2] C. Alexander. Osueke, O. Josphine Ugonabo, T. Williams Sivla, A. John Yakubu, E. Chidinma Okoro, “The First Study on Ionospheric Peak Variability over Equatorial Africa (COSMIC-2),” Advances in Science, Technology and Engineering Systems Journal, 10(4), 14–19, 2025, doi:10.25046/aj100402.

[3] G. Kovács, “Complete System and Interactions of MMF Harmonics in a Squirrel Cage Induction Motor; Differential Leakage; Analytic Calculation,” Advances in Science, Technology and Engineering Systems Journal, 10(4), 20–31, 2025, doi:10.25046/aj100403.

[4] L. Weiguo, C. Yanhong, Y. Libing, Y. Liyan, “Integration and Innovation of a Micro-Topic-Pedagogy Teaching Model under the New Engineering Education Paradigm,” Advances in Science, Technology and Engineering Systems Journal, 10(4), 32–40, 2025, doi:10.25046/aj100404.

Editor-in-chief

Prof. Hamid Mattiello

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Volume 10 Issue 4 **July-August 2025**

CONTENTS

Advances in Science, Technology and Engineering Systems Journal
Vol. 10, No. 4, 1-13 (2025)
www.astesj.com
Special Issue on Computing, Engineering and Multidisciplinary Sciences

ASTES Journal
ISSN: 2415-6698

Machine Learning Methods for University Student Performance Prediction in Basic Skills based on Psychometric Profile

Glender Brás*, Samara Leal, Breno Sousa, Gabriel Paes, Cleberson Junior, João Souza, Rafael Assis, Tamires Marques, Thiago Teles Calazans Silva

UNA University Center, Belo Horizonte, 30640-070, Brazil

ARTICLE INFO

Article history:
Received: 15 March, 2025
Revised: 10 July, 2025
Accepted: 13 July, 2025
Online: 31 July, 2025

Keywords:
Psychometric Profile Assessment
Machine Learning
Student Performance

ABSTRACT

Ensuring the quality of higher education in Brazil presents a complex challenge, intensified by factors that directly affect students' academic performance. The pervasive influence of social media and the overconsumption of superficial digital content undermine students' ability to engage in deep comprehension, critical thinking, and the practical application of knowledge. Furthermore, inadequate preparation during the preceding educational years hinders students' ability to adapt to the academic demands of higher education, leading to difficulties in academic progression and increased dropout rates. In view of the above, this paper explores the potential of Machine Learning models (ML) in predicting the academic performance of higher education students within the Ânima Educação ecosystem, Brazil. The contribution of this work is the development of an artificial intelligence-based assessment tool called AILA that recommends personalized study content for fundamental skills such as Portuguese and Mathematics, based on the psychometric profile of each student. This approach aims to optimize the learning process by addressing individual needs, enhancing academic performance, and overcoming the challenges faced by students in the contemporary educational landscape. Psychometric profile data were collected from approximately 41,296 incoming students of the Ânima Educação universities on the following dimensions: learning, social intelligence, emotional management, socio-emotional skills, teaching method, and knowledge area of the students. The AILA ML models presented good results in predicting students' basic skills performance in the binary and regression approaches. Specifically, the CatBoost model showed an accuracy of 0.74 in predicting scores on the Portuguese and Mathematics and Logical Reasoning proficiency tests.

1. Introduction

This paper is an extension of the paper originally presented at the *2024 IEEE 12th International Conference on Intelligent Systems* [1]. The 2024 Map of Higher Education in Brazil, published by the Semesp Institute, reveals that more than a half of university students (57.2%) drop out before completing their courses [2]. A potential contributor to the high student dropout rate in Brazil's higher education system is the deficient quality of public basic education, largely due to insufficient government investment in the sector. Furthermore, the growing influence of social media and the excessive consumption of superficial digital content have contributed to a reduced capacity for deep comprehension, critical analysis, and the practical application of knowledge. In [3], the authors argued that excessive daily internet use can lead to family conflicts, impaired communication, superficial relationships, learning difficulties, anxiety disorders, and attention deficits. Consequently, students who enter university lacking essential competencies may struggle to maintain academic progress, ultimately leading some to withdraw.

A practical illustration of the substandard quality of education in Brazil is provided by the nation's performance in the Program for International Student Assessment (PISA), a study administered by the OECD (Organization for Economic Cooperation and Development). PISA evaluates the knowledge of 15-year-old students in the domains of reading, mathematics, and science. In the PISA 2022 study, which assessed 81 countries, Brazil ranked among the 20 lowest-performing nations in two of the three evaluated subjects [4]. Moreover, the country scored below the OECD average across all domains. The results indicate that approximately 50% of Brazilian students failed to reach the minimum proficiency level in the assessed areas.

Insufficiency in basic skills represents one of the main challenges faced by university students, directly impacting their academic per-

*Corresponding Author: Glender Brás, UNA University Center, Belo Horizonte, 30640-070, Brazil, glender.medeiros@ulife.com.br

https://dx.doi.org/10.25046/aj100401

formance. According to [5], difficulties in academic writing and text interpretation compromise the ability of students to organize ideas, write coherent texts, and adequately express their knowledge, thus hindering their performance in assessments and academic assignments. Similarly, the study of [6] reveals that gaps in fundamental mathematical skills, such as fractions and problem-solving, can hinder academic progress, especially in courses that require quantitative reasoning. These limitations often necessitate remedial coursework, which extends the duration of studies and contributes to higher dropout rates.

Furthermore, the basic skills play a critical role in the development of higher-order cognitive abilities, such as analytical thinking and problem-solving [7]. The absence of these fundamental skills can significantly impair a student's ability to engage with more complex academic content and tasks. Without a solid foundation in essential areas like literacy and numeracy, students often struggle to make connections between different concepts, analyze problems critically, and apply learned knowledge to real-world situations. As a result, their ability to develop academic autonomy is severely hindered. This lack of autonomy in turn affects their capacity to manage learning independently, which is vital for success in higher education. Students who are unable to independently navigate through academic challenges often rely heavily on external support, such as remedial courses or additional tutoring, which further extends their time in the education system.

In this way, the lack of basic skills not only directly impacts academic performance but also prevents students from developing the necessary self-regulation and problem-solving strategies that are essential for lifelong learning and success in the professional world. Consequently, addressing these deficiencies early in a student's academic career is essential for fostering both academic independence and long-term educational success [8].

Given this, a number of studies have explored the application of machine learning (ML) as a strategy to address the aforementioned issues. Some of the studies involve the use of data analysis and prediction techniques to assess students' academic performance and psychometric profiles. Psychometric profiling is a way to understand an individual's psychological and behavioral characteristics. This is done through psychometric tests that measure the cognitive abilities, personality traits, motivation, interests, and attitudes and can be used to understand how a person thinks, learns, and behaves in different situations [8].

Contemporary psychometrics has shown a significant potential in enhancing psychological assessments, particularly through the integration of ML algorithms [9]. These algorithms possess the ability to improve their performance over time, learning from new data and experiences. This capability allows the inclusion of probabilistic relationships within computer programs, enabling more nuanced and accurate predictions of individual behavior and characteristics. Unlike traditional educational technology tools that rely on predefined, rigidly programmed rules to process inputs and generate outputs, ML models offer a dynamic and adaptive approach. They can continuously evolve based on incoming data, providing greater flexibility and responsiveness in the analysis. This adaptability is a key advantage, as it enables psychometric tools to stay relevant and effective as new patterns and insights emerge, making them far more powerful in real-world applications where data and conditions are constantly changing.

In [10], for instance, a hybrid regression model was proposed to enhance the accuracy of predicting student grades in various subjects. Additionally, an optimized multi-label classifier was developed to qualitatively predict the factors that influence student performance. The model employs three dynamic weighting techniques: collaborative filtering, fuzzy set rules, and Lasso linear regression. This integration of techniques enables a more flexible and adaptable analysis of the variables that impact learning.

Despite significant advances in recent literature on student performance prediction, several limitations persist, including challenges related to imbalanced datasets, unreliable data sources, and concerns regarding the transparency and quality of artificial intelligence (AI) models [11, 12]. These limitations pose substantial challenges in the practical application of predictive tools within real-world educational environments. Specifically, they hinder the ability to generate accurate and reliable insights into student performance, which are essential for formulating evidence-based learning strategies. In the absence of high-quality data and robust model transparency, the reliability of predictions is compromised, making it difficult for educators to make informed decisions. Consequently, this undermines the effectiveness of personalized learning interventions and hampers the creation of adaptive educational strategies that can cater to the diverse needs of students. Overcoming these challenges is crucial for ensuring that predictive analytics can be used to meaningfully enhance educational outcomes and improve the overall learning experience.

In view of the above, this paper presents the AILA (Artificial Intelligence for Learning Assistance) project, that aims to study and implement an AI-based algorithm that utilizes multiple ML models to evaluate students' performances and suggest appropriate academic content to assist the students of Ânima Educação. This tool enhances the accuracy of student performance predictions by integrating diverse data sources, enabling more effective management of unbalanced data sets and improving the transparency of its recommendations.

AILA was developed to enable a more accurate diagnosis of learning gaps, thereby supporting personalized interventions aimed at improving student performance. Leveraging machine learning models, AILA generates individualized learning plans, ensuring targeted support aligned with each student's specific needs. This innovative tool aligns with the institutional and academic objectives of higher education institutions by enhancing retention rates, minimizing the need for remedial instruction, and promoting a more efficient academic trajectory. Ultimately, the algorithm offers a data-driven approach to learning, fostering continuous improvement in student outcomes and advancing educational quality.

The case study of AILA was implemented among incoming students at Ânima Educação Group, with the objective of providing personalized recommendations based on each student's psychometric profile. The Ânima Educação Group is a prominent private educational organization in Brazil, operating 25 educational brands and managing over 500 educational centers nationwide, with a student population of approximately $400,000$.

The data collection process for the mapping phase of this study occurred in two distinct methods. First, a series of questionnaires were answered by university students who have recently entered their courses at one of Ânima's higher education institutions. The Likert scale [13], a prevalent instrument in questionnaire design, was used to employ a five-point scale ranging from "never" to "always." This scale is commonly utilized in the examination of attitudes, beliefs, and behaviors. The questionnaires are organized into three dimensions of knowledge: socio-demographic, socio-cognitive, and socio-emotional. In addition, the students completed Portuguese and Logical Reasoning tests.

These data were pre-processed and, based on the quantitative average of the scores obtained in the questionnaires/tests, the students are mapped as Naive, Beginner, Apprentice or Advanced and receive a recommendation based on this taxonomy. After this mapping, the ML models are then used to predict the students' scores in the Portuguese Language and Logical Reasoning diagnostic tests. This prediction is done considering three approaches:

- **Binary Classification**: The prediction is whether the students achieved high performance (1) or low performance (0) in these tests, based on the average scores of the population and;
- **Multi-class Classification**: The target variable is the classes of the taxonomy.
- **Regression**: The models predict the students' scores;

The case study demonstrates the effectiveness of AILA's machine learning models in predicting academic performance by integrating students' psychometric variables, thereby enhancing predictive accuracy. This approach underscores the importance of incorporating AI into teaching methodologies to support both students and educators. By identifying students' strengths, weaknesses, and potential academic difficulties in advance, AILA enables the provision of targeted resources to mitigate challenges and foster academic success

A total of 41, 296 students completed the aforementioned questionnaires between September 2023 and October 2024 via a custom-developed web application. In addition to data collection, this application features a user-friendly interface that presents the learning content recommended by the models in a clear and accessible manner. The following ML models were employed: CatBoost, Decision Tree (DT), Random Forest (RF), XGBoost, Support Vector Classifier (SVC), and Support Vector Regressor (SVR), all of which demonstrated strong performance across regression, binary, and multi-class classification tasks. For example, the CatBoost model achieved an accuracy of 0.74 in predicting proficiency scores in Portuguese and logical-mathematical assessments using a binary classification approach. In contrast, under a multi-class configuration, the XGBoost and Decision Tree Classifier yielded better results. A comprehensive analysis and discussion of these findings are presented in Section 5.

The rest of this paper is organized as follows: Section 2 presents a summary of the literature, containing works that explore the use of ML models to assess student performance. In Section 3 concepts and methods relevant to this research are discussed. Section 4 explains the processes carried out to test the models and provides a comparative analysis of their performance. Finally, Section 6 presents our conclusions about the study.

2. Related Works

According to [14], predicting students' academic performance has become an increasingly complex task due to the growing volume and variety of data within educational systems. The authors argue that prevailing predictive methods remain insufficient for accurately identifying the most appropriate techniques to assess student performance in higher education institutions. Additionally, the identification of factors influencing student performance remains an underexplored area requiring further investigation, highlighting the need to determine which variables exert the most significant impact on academic outcomes.

Given this, a number of studies have been conducted in the literature to explore the use of ML models as a strategy to assess student performance in various domains. Among these studies, some are particularly noteworthy due to their relevance to the current research and the significant contributions they have made. In [14], the authors conducted a comprehensive review of 162 studies that utilized ML techniques to predict student performance between 2010 and 2022. The study of [15] proposes an intelligent system based on ML to predict students' academic performance, taking into account factors such as attendance, grades, and participation in activities. Algorithms such as Random Forest and Support Vector Machines (SVM), which have been shown to be effective in analyzing academic data, were used. The model developed showed an accuracy of 85%, standing out for its ability to accurately predict performance, with great potential for personalizing pedagogical interventions and optimizing educational outcomes. A key point observed in that research is that the appropriate choice of variables (features) can significantly influence the quality of the predictions.

The use of deep neural networks (DNN) to assess the quality of English language teaching is explored in [16], offering a more effective alternative to traditional methods. With an accuracy rate of 97%, the model is able to process large amounts of data and capture the semantic nuances present in texts, facilitating evaluation in a scalable and less subjective way. The research demonstrates how automating the feature extraction process can reduce cost and time, while improving the accuracy and consistency of scores, bringing an innovative solution to the field of language teaching.

The study of [17] utilizes ML algorithms to identify low-engagement students in a social science course at the Open University (OU) and assess how engagement affects performance. The analysis included variables such as education level, assessment scores, and interactions with virtual learning environment (VLE) activities. Several ML models, including decision trees and gradient-boosted classifiers, were tested, with the best performance in accuracy and recall. A dashboard was developed to help instructors monitor student engagement and provide timely interventions, further exploring the relationship between engagement and course assessment scores.

In [18], the use of ML models is proposed to predict the development of university students' skills over the course of their studies. By analyzing performance data in assessments and extracurricular activities, the authors were able to identify patterns that allow them to predict the evolution of students' cognitive and socio-emotional skills. The results show that deep learning models are effective, achieving 90% accuracy, and provide an agile way to tailor pedagogical approaches to students' individual needs.

The application of ML in the development of flexible learning environments is examined in [19], highlighting its potential to enable new forms of personalized instruction. The study shows how ML models can be used to adapt the content and pace of teaching to the needs of each student, resulting in greater engagement and better academic outcomes. In addition, automating the assessment process helps eliminate human bias, promoting a fairer and more accurate way of assessing students. Based on data collected in real time, the study suggests how curricula and pedagogical strategies can be continuously adjusted to promote more inclusive learning. [20] presents a comprehensive analysis of the literature on how ML has been applied to identify characteristics that affect students' academic performance. The review of 84 publications found that academic and demographic variables, such as grade history and attendance, are the most commonly studied. The study indicates that, although existing models yield satisfactory performance, incorporating additional factors—such as family dynamics and students' psychological characteristics—could enhance predictive accuracy. Moreover, it emphasizes that expanding educational databases is essential to optimizing personalized interventions.

A detailed review of the use of ML in online education is presented in [21], with a particular focus on enhancing student skill acquisition. The study shows that techniques such as content personalization, automatic correction, and progress prediction have been effective in optimizing learning. However, the authors also highlight important challenges, such as privacy issues and model accuracy, and suggest that more research should be done to overcome these limitations. Collaboration between educators, researchers and platform developers is also seen as essential to maximize the positive impact of ML in education. In [22], the authors reviews the main applications of AI and ML in digital education, covering topics such as intelligent tutors, dropout prediction, adaptive learning and process automation. The work shows that artificial neural networks and SVM are the most widely used algorithms, with an emphasis on predictive models aimed at preventing dropout and improving student performance.

The study [23] explores ML application to predict the development of university students' basic skills throughout their course. Using algorithms such as RF and SVM, the research analyzed academic performance data and practical activities to identify the students most likely to succeed or struggle. The ML model proved effective in identifying patterns, allowing for faster and more personalized interventions, which could be crucial in optimizing students' learning and academic development.

In [24], the author focus is on predicting which students are at risk of dropping out of courses on online learning platforms such as MOOCs and Learning Management Systems (LMS). Using machine learning and deep learning algorithms, the study analyzed variables such as performance in assessments, engagement, and online behavior to identify students at risk at an early stage. The model, based on the RF algorithm, achieved excellent results in terms of precision and recall, demonstrating how engagement data can significantly improve the effectiveness of predictions. This allows for timely intervention to prevent students from dropping out and improve their academic performance.

The work of [25] investigates the use of deep neural networks (DNN) to predict the academic performance of students in a data structures course. The model achieved 89% accuracy when using the SMOTE oversampling technique based on students' grades from previous courses. In addition, DNN outperformed other ML algorithms, such as SVM and RFs, on several performance metrics. The research suggests that the model could be a valuable tool for identifying at-risk students early in the semester, enabling early intervention to improve academic outcomes.

Thus, several studies have been proposed to use ML models to improve the design of educational systems and create a more personalized and effective learning experience. However, challenges persist regarding feature selection, dataset size and balance, and the explanatory power and reliability of these models. Furthermore, there is a lack of studies that have applied the results of these models in real-world settings and presented the models' predictions and recommendations through a user-friendly interface.

3. Main Concepts

This section provides an overview of the main concepts and methods used in this study, aiming to make the reading smoother and the understanding more accessible. The goal is to clarify how ML techniques can be applied to assess academic performance, contributing to promoting student success. Section 3.1 outlines the fundamental concepts underlying student performance analysis and their application within the scope of this study. Section 3.2, in turn, presents and defines the ML models employed to predict deficiencies in essential academic skills.

3.1. Elements for Student Performance Analysis

As defined by [26], performance can be understood as the way someone or something acts or behaves, measured by its output. In the educational context, student performance refers to the assessment of students based on criteria that consider essential competencies for the current scenario [27]. In this study, performance analysis is conducted through three main approaches: the Likert Scale [13], the Item Response Theory (IRT) [28], and the principles of Psychometrics [8].

The Likert Scale is a widely used instrument for assessing perceptions and preferences, and it is classified as a summative assessment method [13]. It offers response options arranged in a progressive order, typically ranging from strong disagreement to strong agreement. In the context of this study, the Likert Scale is employed to measure an individual's self-perceived proficiency in a given skill, using a five-point scale with the following categories: "never," "rarely," "sometimes," "often," and "always".

The concept of psychometrics can be approached in various ways, one of which is its application in assessing an individual's psychological traits [29]. In this sense, psychometrics involves the development of measurement tools, such as tests, scales, and questionnaires, to perform a precise and valid analysis of different aspects of human behavior [8]. Moreover, psychometrics extends beyond simple measurement by emphasizing the quality and accuracy of assessment instruments. Its primary aim is to ensure that these instruments yield consistent results while effectively capturing the constructs they are intended to measure [30]. Guided by these principles, this study applied psychometric concepts to construct the profile of the analyzed students, as detailed in Section 4.

3.2. Machine Learning (ML)

The field of machine learning focuses on developing algorithms capable of learning directly from data rather than following predefined commands [31]. The primary goal is to build computational systems that, when fed with a dataset, can generate models to make predictions, classifications, or identifications based on the acquired knowledge. Within this context, we applied various ML approaches to address the proposed problem, including Decision Tree, Random Forest, Neural Network, and Support Vector Machine. The ML techniques used in this study were:

- **Decision Tree (DT):** According to [32], a Decision Tree [33] is a hierarchical, branched structure composed of nodes and branches. At each internal node, a decision is made based on a test applied to input variables, guiding the flow along specific branches. The terminal nodes, or leaf nodes, provide the predicted values of the target variable or the associated probability distributions. According to [34], the Decision Tree, known for its quick understanding and ease of implementation, is often adopted in decision support systems in the healthcare field. Its versatility makes it applicable in various domains, offering benefits in terms of efficiency and simplified operation. With the ability to generate clear and easy-to-understand analyses, DT establishes itself as a valuable tool to optimize decision-making processes in different contexts, standing out for its accessibility and effectiveness in various areas of knowledge.

- **Random Forest (RF):** The Random Forest method, as described by [33], consists of a set of classifications based on decision trees, where each tree is influenced by a random vector. This vector is generated independently and follows a uniform distribution among the trees in the forest. Various approaches can be applied to construct these vectors, including bagging, estimated selection of splits, output randomization, and estimated attribute selection. The fundamental principle of this model lies in the independence of the generated vectors for each tree individually. By gathering a large number of trees and combining their decisions, the technique aims to increase accuracy in data classification for ML problems. This is achieved because, after building the trees, the final prediction is determined based on the most voted class by the ensemble [35].

- **Support Vector Machine (SVM):** Introduced by [36], SVM is a ML technique for binary classification problems. Its approach involves transforming input vectors through a non-linear mapping into a high-dimensional space, where a linear decision boundary is constructed with specific characteristics that ensure good generalization capability. Originally, this model was developed to handle perfectly separable datasets but was later improved to accommodate scenarios where data exhibits overlap and is not completely separable.

- **XGBoost:** As detailed by [37], XGBoost is a ML model based on decision trees, designed to maximize computational efficiency and predictive accuracy. This method is distinguished by using "boosting," an approach where each subsequent tree seeks to correct the errors of the previous ones. XGBoost differs from conventional methods by integrating regularization mechanisms that minimize the risk of overfitting, applying, among other strategies, data partitioning and parallelism in tree construction. Its ability to scale processing of large data volumes makes it a popular choice for a wide range of predictive problems. Additionally, the model enhances classification accuracy by coordinating the integration of multiple trees, focusing on reducing bias and variance. According to [37], XGBoost was developed to be an effective, versatile, and easy-to-implement tool, making it a predominant choice in competitions and practical applications.

- **CatBoost:** It is an ML algorithm that stands out for its ability to optimize performance in decision tree-based models, especially when dealing with categorical data. The algorithm's main innovation is the use of Ordered Target Statistics, a method that improves the encoding of categorical variables and thus reduces the risk of overfitting by preventing information from leaking improperly during training [38]. Additionally, CatBoost implements a symmetric boosting approach, which ensures greater training efficiency and enhances the model's ability to generalize to new data. These advancements make CatBoost an effective solution for classification and regression problems, particularly useful in contexts with large data volumes and challenging tasks, such as those in finance and marketing sectors [39].

4. Proposed Modeling

In this section, we present the models implemented for understanding the relationship between the psychometric profiles of the students and their performance in basic skills. The primary objective of this study is to develop the ML models, which aim to assess both the knowledge of the Portuguese Language and Mathematical Reasoning. The implementation of these methods will facilitate the determination of the relationship between different psychometric domains and the teaching and learning of these fundamental subjects, which can support the students with the recommendation of relevant content to improve their abilities. The information was modeled based on three output configurations:

- **Binary Classification**: categorizes students into two groups: "high performance" (1) or "low performance" (0).

- **Multi-class classification**: classifies students based on the four classes or taxonomies: Naive, Beginner, Apprentice, and Advanced, with Naive having the lowest scores and Advanced the highest.

- **Regression**: the regression analysis, in turn, is used to predict the scores in the assessments, showing the improvement of the students. In addition, logistic regression, based on the idea of 'propensity scores', is used to identify the correlation between student grades and other factors, helping to understand the elements that influence their grades [40].

The development of the predictive models was guided by the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology [41], ensuring a structured, systematic, and replicable approach. The use of CRISP-DM enabled a logical progression from data understanding to model implementation, ensuring that each step effectively contributed to the quality and accuracy of the predictions. Figure 1 shows the process of this methodology.

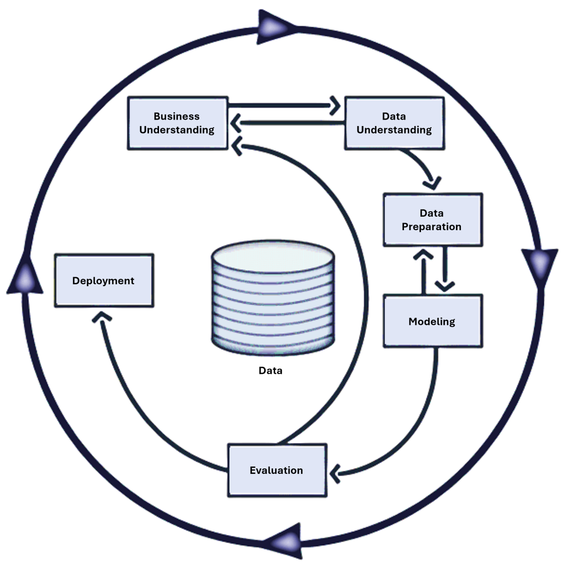

Figure 1: Phases of CRISP-DM applied to this study.

The application of CRISP-DM began with **business and data understanding**, during which the project objectives were defined and the characteristics of the data collected from student questionnaires were analyzed. These questionnaires — comprising emotional states, cognitive behaviors, and academic routines — underwent a rigorous preparation process conducted by educators from Ânima, which was essential to ensuring data quality and reliability.

During the **data preparation** phase, standardization and encoding techniques were applied to ensure data consistency and compatibility, optimizing it for subsequent modeling. This careful preparation was essential for generating a robust dataset that served as the foundation for building effective predictive models.

Following the data preparation stage, the process advanced to the **modeling** and **evaluation** phases, during the ML models were trained, fine-tuned, and rigorously assessed for performance. Finally, in the **deployment** phase, AILA's models and their artifacts were structured to ensure reusability and seamless integration into production environments. A web-based application was developed to make the models' personalized learning recommendations accessible to students in a user-friendly and practical manner, thereby completing the CRISP-DM cycle and establishing this study as a practical contribution to real-world educational contexts.

4.1. Data Understanding and Preparation

As previously mentioned, the data used to train and test the ML models were collected through psychometric questionnaires designed to assess students' emotional states, cognitive behaviors, and academic routines, along with diagnostic tests measuring proficiency in fundamental skills (Portuguese and Logical Reasoning). These instruments were administered to incoming students at Ânima Educação institutions, as illustrated in Figure 2.

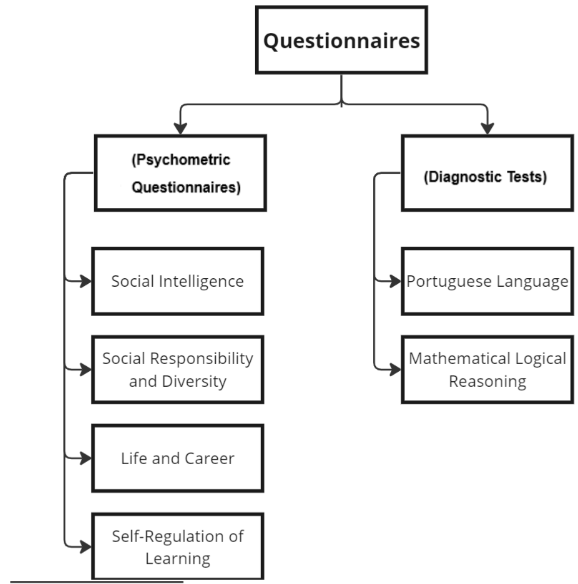

Figure 2: Questionnaire Flowchart

It is important to note that the aforementioned data were processed in accordance with Brazil's General Data Protection Law (Lei Geral de Proteção de Dados Pessoais – LGPD), which safeguards privacy and ensures information security. Participation in the study was limited to students who provided full consent, reinforcing ethical research practices and legal compliance established by Ânima Educação institutions. Although participation was voluntary, several measures were taken to mitigate self-report bias and preserve the diversity and representativeness of the sample. These included the use of validated psychometric instruments with clear, neutrally worded items; the assurance of anonymity and confidentiality to reduce social desirability effects; the incorporation of consistency checks across similar items; and the inclusion of control questions to identify inattentive responses. Moreover, behavioral frequency-based questions were prioritized over abstract self-assessments. Prior to responding, students were clearly informed about the purpose of the study and were encouraged to answer honestly

All data collected in this study were securely stored on AWS infrastructure, through an account managed by Ânima. Access to both the source code and the MongoDB database was strictly limited to the project team, in compliance with privacy and data protection requirements. The application and database are hosted on an EC2 instance, ensuring controlled and secure access aligned with industry-standard data security practices.

The psychometric questionnaires were constructed using the Likert Scale, discussed in Section 3.1, and subjected to a statistical

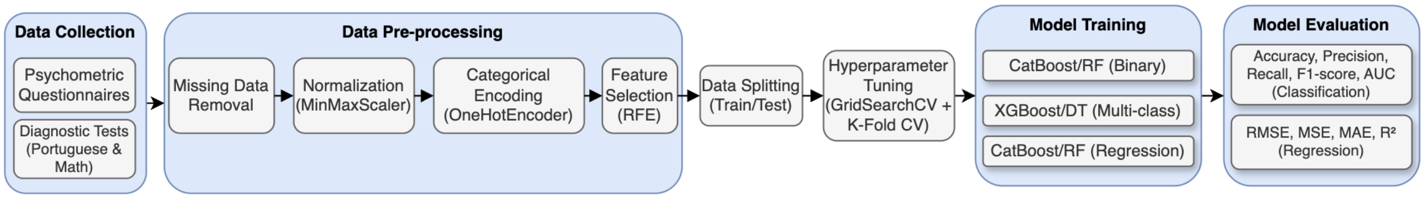

Figure 3: ML Workflow

analysis to aim to validate their effectiveness in mapping the correct profiles based on population analysis. The diagnostics were developed by the specialized professionals in pedagogical intervention.

4.2. Modeling

Figure 3 provides an overview of the ML workflow designed for this study. This modeling framework encapsulates the main stages of the process, from the collection of psychometric and diagnostic data to the training and evaluation of predictive models. This representation provides a clear and systematic overview of the process, enhancing methodological transparency and supporting the reproducibility of the study.

4.2.1. Data Pre-processing

The preliminary processing of data is of the utmost importance for the construction of predictive models, as it ensures the quality of the information. To this end, we employed data derived from questionnaires described in Section 4.1. To ensure data standardization and compatibility, we implemented variable normalization and coding techniques.

We applied the numerical data to undergo *MinMaxScaler* normalization, which scales the values between 0 and 1. This normalization enhances model stability by mitigating distortions caused by features with varying scales and units, such as the number of clicks on VLE activities compared to assessment scores [42]. By standardizing the numerical data, the model can more effectively learn relevant patterns without being disproportionately influenced by any single feature.

For categorical variables, the *OneHotEncoder* method was used, converting each category into binary representations. This approach ensures that the model does not assign any hierarchical or ordinal relationships between categories, treating each one independently and without bias [42]. By encoding variables such as different VLE activities (e.g., forums, resources) separately, the model can better capture the impact of each activity on student engagement. These preprocessing techniques enhance the quality of the data and improve the ability of the ML models to accurately predict student engagement.

4.2.2. Used Models

A comparative analysis was conducted to evaluate the efficacy of several models to predict student's performance in this study. The analysis revealed that XGBoost exhibited superior performance in terms of efficiency in decision trees and the capacity to process substantial volumes of data. CatBoost also demonstrated satisfactory results, particularly in scenarios involving disorganized data. Additionally, simple decision trees were utilized for comparative purposes, along with Random Forest, which was noteworthy for its stability and predictive capacity. For complex data, SVM was used, as it performs well on many variables [43].

4.2.3. Parameter Adjustment and Validation

The models were adjusted using the method *GridSearchCV*, an exhaustive search of various combinations of adjustments, with the aim of identifying the best option for prediction. To prevent the model from learning too much from the training data alone, k-fold cross-validation was used, ensuring a good forecast on new data. Studies demonstrate that this practice improves the prediction of academic performance by reducing statistical errors [44].

4.2.4. Evaluation Metrics

Accuracy is a widely used metric for evaluating the performance of a specific model, reflecting the ratio of correct predictions to total observations [45]. This metric can be used to further evaluate a model by measuring the ratio of correctly predicted positive cases to total predicted positive cases. This metric is advantageous in situations with high costs of false positive results [46].

The classification and regression methods were assessed using various methodologies, enabling a more comprehensive analysis [47]. Accuracy, which represents "hits," performs optimally with balanced data, while the area under the ROC curve (AUC) is more suitable for unbalanced data [45]. Additionally, the accuracy of positive predictions and the hit rate on positives were evaluated, which are fundamental in academic settings [46]. In the context of regression models, the mean squared error (MSE) was employed as a simplification, prioritizing significant errors while making predictions in a linear fashion [48]. The root mean squared error (RMSE) is the average error of the model expressed in the variable.

5. Case Study

This section details the case study developed to test the validity of the proposed methodologies, contextualizing the study and its results. The study involved 41,296 incoming students from various Ânima Educação institutions. The primary objective was to assist students in overcoming their challenges from the very beginning of their higher education journey, thereby optimizing their academic trajectory right from the outset. To achieve this, the selection of the student sample for the algorithms was based on the availability of complete individual data. Specifically, only students with comprehensive information from all the applied tests were included in the sample.

The distribution of binary classes reveals a significant imbalance in both Portuguese and Math scores. In both subjects, the vast majority of students scored below the classification threshold, while a considerably smaller fraction achieved or exceeded this mark. This class imbalance demands attention in the development of predictive models, as it can negatively impact model performance and generalization capabilities (Figure 4).

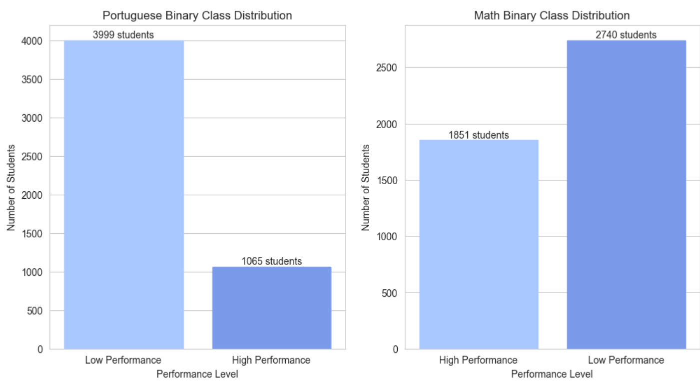

Figure 4: Class distribution Binary Classification

Furthermore, the distribution of learning taxonomies presents interesting patterns. In Portuguese, the "Beginner" category shows the highest concentration of students, while the "Advanced" category is the least represented. In Math, a similar pattern is observed, with the "Beginner" category predominating and "Advanced" being the least common. This uneven distribution among learning categories suggests a need for differentiated pedagogical approaches to address the specific needs of each group (Figure 5).

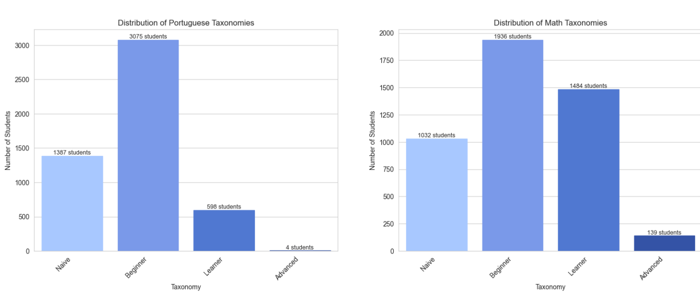

Figure 5: Class distribution Multi-class Classification

This approach was essential, as the supervised learning algorithms used in the study require a full set of labeled data to effectively learn and make accurate predictions. Incomplete data could lead to biased models or reduced prediction accuracy, making it crucial to ensure that only students with complete data were considered. By focusing on students with full datasets, the study aimed to maximize the reliability and validity of the predictions generated by the ML algorithms.

5.1. Experiments

To carry out the experiments, a process was structured that involved analyzing the students' answers to a series of questions, with the aim of measuring different aspects of learning. From these answers, a consolidated data set was generated in a CSV file, containing the features shown in Table 1. The implementation of the AILA algorithm was carried out in the Visual Studio Code environment, using the Jupyter extension to facilitate interactive code execution. The language chosen was Python, due to its wide range of specialized ML libraries.

Table 1: Features and Outputs

Feature	Output
Modality	Logical/Math Reasoning Tax.
Knowledge Area	Portuguese Lang. Tax.
Tax. Learning	Score Logical/Math Reasoning
Tax. Soc. Intell.	Score Portuguese Lang.
Tax. Life/Career	-
Tax. Emot. Mgmt.	-
Tax. Soc. Resp.	-
Score Learning	-
Score Soc. Intell.	-
Score Life/Career	-
Score Emot. Mgmt.	-
Score Soc. Resp.	-

Initially, the data was subjected to a preparation pipeline, which included removing missing values and transforming the variables. To avoid bias in the models, incomplete entries were eliminated, resulting in 4,499 lines. Subsequently, the numerical attributes were normalized using *MinMaxScaler*, ensuring that all the variables were on the same scale, in the 0 to 1 range. *OneHotEncoder* was used to transform categorical variables into numerical representations suitable for machine learning algorithms. The least relevant variables, based on their predictive importance, were gradually removed by Recursive Feature Elimination (RFE), reducing the dimensionality of the data set and improving the performance of the models.

This study used the approach of splitting the data into training and test sets with the help of scikit-learn's train-test-split function. This technique makes it possible to split the data randomly, so that a fraction of it is used to train the model, while the other fraction is used to evaluate its performance.

After preparing the data, the model's hyperparameters were optimized using *GridSearchCV*, a method that systematically goes through a grid of predefined values to find the best combination of parameters. In the experiment, cross-validation was applied with three divisions (cv=3), and the accuracy metric was used as the evaluation criterion. The optimized set of hyperparameters was then selected, and the final model was adjusted based on this configuration, ensuring better predictive performance.

5.2. Results

The experiments were carried out with the aim of predicting the academic performance of the university's students and showed that the models based on boosted decision trees (Boosting) and Random Forest obtained the best results, especially in binary classification. The selection of attributes included variables related to the dimensions of learning, social intelligence, emotional management, and other socio-emotional skills, as well as the teaching method and area of knowledge of the students.

The study also highlighted the influence of teaching methods and students' subject knowledge on academic performance. Certain pedagogical strategies appeared to promote better learning outcomes, while discipline-specific factors also played a role in shaping students' outcomes. These findings suggest that personalized interventions, tailored to students' academic and socio-emotional profiles, could be instrumental in improving educational outcomes. Future research could explore the impact of these interventions and further refine predictive models to support data-driven decision-making in educational settings.

5.2.1. Binary Classification

The *CatBoostClassifier* and *RandomForestClassifier* models performed better in predicting students' proficiency in Portuguese and logical and mathematical reasoning. In predicting Portuguese, the CatBoost model obtained the best results, while for logical and mathematical reasoning, the *RandomForest* model performed better, as shown in Table 2.

In addition, it was observed that the models showed high metrics for class 0, which encompasses the students with the greatest difficulty in the subjects assessed. This result is particularly relevant, as it indicates that the models are able to more accurately identify the students who need the most attention and pedagogical support. The high precision and recall for class 0 reinforce the ability of these approaches to correctly discriminate between students with difficulties, making them useful tools for targeted educational interventions.

The joint analysis of the confusion matrices from the *CatBoostClassifier*, shown in Figure 6, complements these findings. For Portuguese Language, there is a good performance in identifying students in class 0, but lower accuracy in identifying those in class 1. In the case of Logical and Mathematical Reasoning, the model produced more balanced results between the classes. This combined visualization highlights the models' focus on correctly identifying students with low performance, which is the central objective of this study.

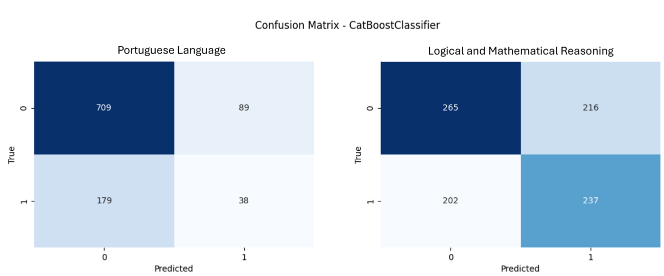

Figure 6: Confusion Matrix - CatBoostClassifier

The superior performance of the Boosting and Random Forest models can be justified by the fact that these techniques employ ensemble learning, reducing bias and variance [49]. CatBoost, in particular, uses effective processing of categorical variables and reduces the impact of overfitting, while Random Forest benefits from the aggregation of multiple decision trees, promoting robustness to the model [38].

Table 2: Binary Classification Results. PT = Portuguese language; LR = Logical and Mathematical Reasoning. Acc = Accuracy; Prec 0 = Precision for class 0; Prec 1 = Precision for class 1; F1 = F1-Score (weighted).

Model	Type	Acc	Prec 0	Prec 1	F1
CatBoost	PT	0,74	0,80	0,30	0,22
XGB	PT	0,71	0,80	0,26	0,22
DecisionTree	PT	0,53	0,82	0,25	0,35
RandomForest	PT	0,71	0,80	0,26	0,28
SVC	PT	0,53	0,82	0,25	0,35
CatBoost	LR	0,55	0,57	0,52	0,53
XGB	LR	0,52	0,55	0,50	0,51
DecisionTree	LR	0,56	0,57	0,54	0,53
RandomForest	LR	0,56	0,57	0,54	0,53
SVC	LR	0,55	0,56	0,53	0,51

5.2.2. Multi-class Classification

With regard to multi-class classification, the XGBoost and *DecisionTreeClassifier* models showed the best results for Portuguese language and logical and mathematical reasoning, respectively (Table 3). These results suggest that although Boosting models remain effective, the complexity of predicting multiple classes may have affected overall accuracy.

Table 3: Multi-class Classification Results. PT = Portuguese language; LR = Logical and Mathematical Reasoning. Acc = Accuracy; Prec = Macro-averaged Precision; F1 = Macro-averaged F1-Score.

Model	Type	Acc	Prec	F1
CatBoost	PT	0,47	0,48	0,47
XGB	PT	0,56	0,48	0,49
DecisionTree	PT	0,56	0,49	0,51
RandomForest	PT	0,43	0,48	0,45
SVC	PT	0,35	0,47	0,38
CatBoost	LR	0,37	0,36	0,36
XGB	LR	0,40	0,38	0,38
DecisionTree	LR	0,42	0,40	0,39
RandomForest	LR	0,37	0,37	0,37
SVC	LR	0,35	0,37	0,35

As shown in Figure 5, the used dataset has a much larger number of students with low performance than students with high performance. This can bias the model's learning, especially when it involves multi-class classification, and make it difficult to identify different levels among the students. For this reason the multi-class models found a low accuracy, as shown in Table 3.

Despite these limitations, the multi-class classification models still provide complementary insights regarding the distribution of student performance levels. The scarcity of examples in intermediate and high-performance classes hinders the models' ability to learn discriminative patterns for these categories, contributing to the overall lower accuracy. For the model to better classify student performance levels through the multi-class approach, more data would be needed in the other classes. However, since the critical task of identifying students with unsatisfactory performance is handled by the binary classification model—more robust and less affected by class imbalance—the limitations of the multi-class model do not compromise the practical objectives of this study. Thus, multi-class results should be viewed as a secondary analytical tool, useful for exploratory interpretation, while pedagogical decision-making is grounded on the binary model's outputs.

5.2.3. Regression

Although the *CatBoostRegressor* and *XGBRegressor* achieved the best results for Portuguese, and the *RandomForestRegressor* for logical and mathematical reasoning (Table 4), the regression models generally exhibited low predictive power. This limitation may be related to the fact that the domain scores used as predictors are not necessarily strong indicators of performance in another domain, as each one assesses distinct cognitive skills. Therefore, attempting to predict performance in a specific area based on performance in others may fail to adequately capture the unique characteristics of each evaluated competency.

Table 4: Regression Model Performance. PT = Portuguese language; LR = Logical and Mathematical Reasoning. RMSE = Root Mean Squared Error; MSE = Mean Squared Error; MAE = Mean Absolute Error; R² = Coefficient of Determination.

Model	Type	RMSE	MSE	MAE	R^2
CatBoost	PT	0.16	0.03	0.13	0.05
XGB	PT	0.16	0.03	0.14	0.04
DecisionTree	PT	0.17	0.03	0.14	-0.02
RandomForest	PT	0.16	0.03	0.14	0.04
SVR	PT	0.17	0.03	0.13	0.03
CatBoost	LR	0.25	0.06	0.22	0.01
XGB	LR	0.25	0.06	0.22	0.00
DecisionTree	LR	0.25	0.06	0.22	-0.04
RandomForest	LR	0.25	0.06	0.22	0.01
SVR	LR	0.25	0.06	0.22	-0.02

5.3. Discussion of the Results

The results obtained confirm that the Boosting and Random Forest models are highly effective for binary classification problems, corroborating previous studies that highlight their ability to capture complex patterns and reduce overfitting through regularization [50] [37].

The analysis of variable importance in the CatBoost model, used to predict students' performance in the Portuguese language, reveals which features had the greatest influence on the predictions. Figure 7 presents a bar chart ranking the most relevant features based on their importance values.

It is observed that scores related to life and career and learning had the greatest impact, followed by social intelligence and emotional management. These results suggest that socioemotional skills play a crucial role in students' performance, not only in content mastery but also in their ability to apply such knowledge in the exam.

Additionally, categorical variables such as the field of study (Humanities, Biological and Health Sciences, among others) and the mode of instruction (in-person or not) also showed some influence, although to a lesser extent, in predicting the results. This analysis reinforces the need for educational policies that consider not only technical knowledge but also interpersonal and emotional skills, which directly impact students' academic success in the Portuguese language exam.

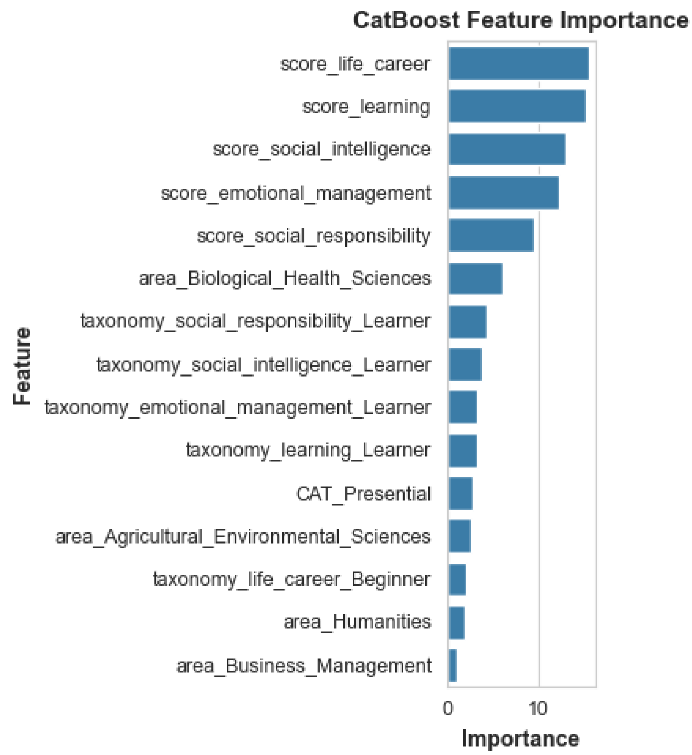

Figure 7: Feature Importances – CatBoost Model

However, multiclass prediction proved to be more challenging, possibly due to the imbalance in the data, which can guide the models learning to a good accuracy in one class but not in the others. Regarding regression, the relatively low R^2 values suggest that other factors, such as individual aspects of the students and unmeasured external factors, may have influenced academic performance.

5.4. Practical application of the recommendation

AILA employs the responses of students to psychometric questionnaires and diagnostic tests to categorize them according to a taxonomy comprising four levels: The designation of Naive, Beginner, Apprentice or Advanced is utilized to categorize individuals based on their level of expertise or proficiency in a particular domain. In accordance with the classification that has been presented, the platform provides recommendations that are customized to align with each student's unique profile.

Figure 8 presents a simulated example of the AILA's recommendation interface, which displays content suggestions categorized by dimension (e.g., logical reasoning, social intelligence, life and career). The recommendations presented are based on the student's level in each domain, with the objective of strengthening specific skills.

The practical implementation in question establishes a connection between psychometric assessment and pedagogical guidance, promoting a more student-centered learning experience.

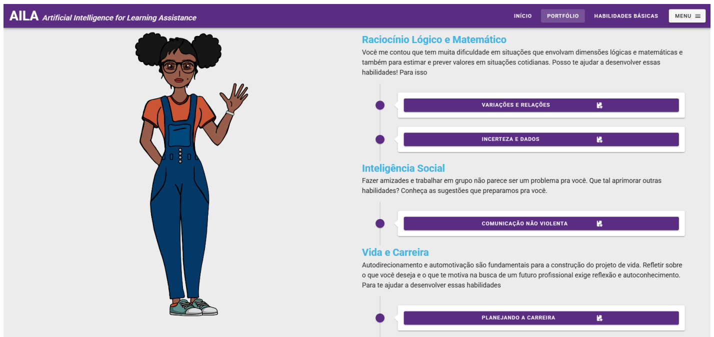

Figure 8: Simulated interface displaying personalized recommendations generated by AILA based on psychometric classification

6. Conclusion

This work explored ML models to university students performance in basic skills such as Portuguese language and Logical Reasoning. The study has been applied in a real case involving more than $40,000$ students at several institutions of Ânima Educação, an educational group in Brazil.

The results reinforce the effectiveness of using ML to predict the academic performance of higher education students and show that the models implemented are effective for predicting student performance in basic skills, especially for predicting students who will have difficulty, which is the main objective of this study. The high accuracy in class zero, in the case of binary classification, reinforces this point and can provide support to intervention actions on the part of the university. Similarly, the regression models demonstrated effectiveness in predicting the student's score and can also serve as an important method for this purpose.

As for the multi-class classification task, the models presented a low accuracy and showed that the proposed models are still not sufficient to find several levels among the students satisfactory. This is justified by the imbalance of data in the used dataset and the low amount of data in classes that represent high student performance, that is, the low number of students with high performance in basic skills in the research carried out. This confirms that data balance is crucial for this type of task.

In addition, the models proved to be effective in finding a correlation between psychometric profile and performance in basic subjects, which suggests that emotional and psychosocial factors can influence the learning process of students. For future work, it will be important to conduct new tests with more advanced students in order to improve the performance of the multi-class model, thus generating advances in the learning of the model for these cases.

However, a central limitation of this study lies in the pronounced imbalance in class distribution, particularly in the context of multi-class classification. The scarcity of students with high performance in both Portuguese Language and Logical-Mathematical Reasoning restricted the models' ability to learn representative patterns across all performance levels. This imbalance contributed to the lower accuracy observed in the multi-class models and compromised the regression results by reducing the diversity and richness of the training data. Moreover, the reliance on correlated data, such as psychometric profiles based on self-reported responses, may introduce potential bias, as subjective answers can reflect inconsistent or distorted perceptions. Future work should address these limitations through the collection of more balanced datasets, the application of techniques to mitigate self-report bias, and the use of data augmentation and resampling methods, aiming to improve the models' generalization capacity and robustness.

References

[1] G. Brás, S. S. Leal, B. Sousa, C. Junior, J. Souza, Z. Mendes, G. Paes, "Machine Learning Models for Basic Skills Identification in University Students," in 2024 IEEE 12th International Conference on Intelligent Systems (IS), 1–9, IEEE, 2024, doi:10.1109/IS61756.2024.10705207.

[2] I. Semesp, "Mapa do Ensino Superior no Brasil," in 14ª Edição, 39, 2024.

[3] T. d. O. Silva, Os impactos sociais, cognitivos e afetivos sobre a geração de adolescentes conectados às tecnologias digitais, Undergraduate thesis (tcc), Universidade Federal da Paraíba, João Pessoa, Brazil, 2016, submitted June 14, 2016; bibliographic study; advisor: Lebiam Tamar Gomes Silva.

[4] D. Salinas, F. Avvisati, R. Castaneda Velle, "PISA 2022 Results: The State of Learning and Equity in Education – Volume I," Technical report, Organisation for Economic Co-operation and Development (OECD), Paris, France, 2023, doi:10.1787/53f23881-en.

[5] R. Bailey, "Student writing and academic literacy development at university," Journal of Learning and Student Experience, **1**, 7–7, 2018.

[6] F. Ngo, "Fractions in college: How basic math remediation impacts community college students," Research in Higher Education, **60**, 485–520, 2019, doi:10.1007/s11162-018-9519-x.

[7] L. Aranda, E. Mena-Rodríguez, L. Rubio, "Basic skills in higher education: An analysis of attributed importance," Frontiers in Psychology, **13**, 752248, 2022, doi:10.3389/fpsyg.2022.752248.

[8] G. Kuan, Y. C. Kueh, N. Abdullah, E. L. M. Tai, "Psychometric properties of the health-promoting lifestyle profile II: cross-cultural validation of the Malay language version," BMC public health, **19**(1), 1–10, 2019, doi:10.1186/s12889-019-7109-2.

[9] V. R. Franco, "Aprendizado de Máquina e Psicometria: Inovações Analíticas na Avaliação Psicológica," PePisic Periodicos de Psicologia, **20**, a.c, 2021, doi:10.15689/ap.2021.2003.ed.

[10] A. Alshanqiti, A. Namoun, "Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification," IEEE Access, **8**, 203827–203844, 2020, doi:10.1109/ACCESS.2020.3036572.

[11] E. P. F. Kelvin dos Santos, "A Previsão de Evasão em Cursos de Graduação Utilizando Machine Learning," Caderno Virtual, **1**(58), 13–15, 2024.

[12] A. da Silva Franqueira, B. F. C. ZanettiCarlos, A. L. Bitencourt, D. Z. Franco, E. H. B. de Oliveira, Érica Rafaela dos Santos Campos, H. G. M. Júnior, J. da Cruz Chagas, "Análise impulsionada por IA para previsão de desempenho estudantil," Cuadernos de Educación y Desarrollo, **16**(5), 14–17, 2024.

[13] A. M. Feijó, E. F. R. Vicente, S. M. Petri, "O uso das escalas Likert nas pesquisas de contabilidade," Revista Gestão Organizacional, **13**(1), 27–41, 2020, doi:10.22277/rgo.v13i1.5112.

[14] A. Abu Saa, M. Al-Emran, K. Shaalan, "Factors affecting students' performance in higher education: a systematic review of predictive data mining techniques," Technology, Knowledge and Learning, **24**(4), 567–598, 2019, doi:10.1007/s10758-019-09408-7.

[15] D. Petkovic, K. Okada, M. Sosnick, A. Iyer, S. Zhu, R. Todtenhoefer, S. Huang, "Work in progress: a machine learning approach for assessment and prediction of teamwork effectiveness in software engineering education," in 2012 frontiers in education conference proceedings, 1–3, IEEE, 2012, doi:10.1109/FIE.2012.6462205.

[16] J. Zhu, C. Zhu, Z. Wang, "Application of Machine Learning in English Language Teaching Quality Assessment," in 2024 International Conference on Interactive Intelligent Systems and Techniques (IIST), 300–304, IEEE, 2024, doi:10.1109/IIST62526.2024.00041.

[17] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, "Student engagement predictions in an e-learning System and their impact on student course assessment scores," Computational intelligence and neuroscience, **2018**(1), 6347186, 2018, doi:10.1155/2018/6347186.

[18] J. G. Valen-Dacanay, T. D. Palaoag, "Exploring The Learning Analytics Of Skill-Based Course Using Machine Learning Classification Models," in 2023 11th International Conference on Information and Education Technology (ICIET), 411–415, IEEE, 2023, doi:10.1109/ICIET56899.2023.10111210.

[19] A. Ravuri, M. Lourens, S. Aswini, G. Nijhawan, R. S. Zabibah, R. Chandrashekar, "Improving Personalized Education: A Machine Learning Method for Flexible Learning Environments," in 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), 1715–1720, IEEE, 2023, doi:10.1109/UPCON59197.2023.10434888.

[20] I. Issah, O. Appiah, P. Appiahene, F. Inusah, "A systematic review of the literature on machine learning application of determining the attributes influencing academic performance," Decision analytics journal, **7**, 100204, 2023, doi:10.1016/j.dajour.2023.100204.

[21] Y. Christian, Y. Choo, N. Yusof, "Systematic literature review on the use of machine learning in online learning in the context of skill achievement," Journal of Theoretical and Applied Information Technology, **102**(6), 2466–2479, 2024.

[22] H. Munir, B. Vogel, A. Jacobsson, "Artificial intelligence and machine learning approaches in digital education: A systematic revision," Information, **13**(4), 203, 2022, doi:10.3390/info13040203.

[23] D. Petkovic, M. Sosnick-Pérez, S. Huang, R. Todtenhoefer, K. Okada, S. Arora, R. Sreenivasen, L. Flores, S. Dubey, "Setap: Software engineering teamwork assessment and prediction using machine learning," in 2014 IEEE frontiers in education conference (FIE) proceedings, 1–8, IEEE, 2014, doi:10.1109/FIE.2014.7044199.

[24] M. Adnan, A. Habib, J. Ashraf, S. Mussadiq, A. A. Raza, M. Abid, M. Bashir, S. U. Khan, "Predicting at-risk students at different percentages of course length for early intervention using machine learning models," Ieee Access, **9**, 7519–7539, 2021, doi:10.1109/ACCESS.2021.3049446.

[25] A. Nabil, M. Seyam, A. Abou-Elfetouh, "Prediction of students' academic performance based on courses' grades using deep neural networks," IEEE Access, **9**, 140731–140746, 2021, doi:10.1109/ACCESS.2021.3119596.

[26] A. Namoun, A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," Applied Sciences, **11**(1), 237, 2020, doi:10.3390/app11010237.

[27] A. M. J. d. Andrade, Desempenho acadêmico, permanência e desenvolvimento psicossocial de universitários: relação com indicadores da assistência estudantil, Master's thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, 2014, doi:10.1590/S1414-40772017000200014.

[28] E. A. C. d. Araujo, D. F. d. Andrade, S. L. V. Bortolotti, "Teoria da resposta ao item," Revista da Escola de Enfermagem da USP, **43**, 1000–1008, 2009, doi:10.1590/S0080-62342009000500003.

[29] R. Primi, "Psicometria: fundamentos matemáticos da Teoria Clássica dos Testes," Avaliação Psicológica, **11**(2), 297–307, 2012.

[30] C. F. Collares, W. L. P. Grec, J. L. M. Machado, "Psicometria na garantia de qualidade da educação médica: conceitos e aplicações," Science in Health, **3**(1), 33–49, 2012.

[31] G. M. d. M. Paixão, B. C. Santos, R. M. d. Araujo, M. H. Ribeiro, J. L. d. Moraes, A. L. Ribeiro, "Machine learning na medicina: revisão e aplicabilidade," Arquivos brasileiros de cardiologia, **118**, 95–102, 2022, doi:10.36660/abc.20200596.

[32] C. A. Meira, L. H. Rodrigues, S. A. Moraes, "Análise da epidemia da ferrugem do cafeeiro com árvore de decisão," Tropical Plant Pathology, **33**, 114–124, 2008, doi:10.1590/S1982-56762008000200005.

[33] L. Breiman, "Random forests," Machine learning, **45**, 5–32, 2001.

[34] J. P. da Silva Funchal, D. F. Adanatti, "Um Estudo Sobre a Classificação de Risco na Área da Saúde Utilizando Árvores de Decisão," iSys-Brazilian Journal of Information Systems, **9**(3), 89–111, 2016.

[35] T. M. Oshiro, Uma abordagem para a construção de uma única árvore a partir de uma Random Forest para classificação de bases de expressão gênica, Ph.D. thesis, Universidade de São Paulo, 2013.

[36] C. Cortes, V. Vapnik, "Support-vector networks," Machine learning, **20**, 273–297, 1995.

[37] T. Chen, C. Guestrin, "Xgboost: A scalable tree boosting system," in Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785–794, 2016, doi:10.1145/2939672.2939785.

[38] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, A. Gulin, "CatBoost: unbiased boosting with categorical features," Advances in neural information processing systems, **31**, 2018.

[39] J. T. Hancock, T. M. Khoshgoftaar, "CatBoost for big data: an interdisciplinary review," Journal of big data, **7**(1), 94, 2020, doi:10.1186/s40537-020-00369-8.

[40] P. J. L. Adeodato, "Data Mining Solution for Assessing Brazilian Secondary School Quality Based on ENEM and Census Data," 13th CONTECSI International Conference on Information Systems and Technology Management, online, 2016.

[41] J. L. C. Ramos, R. L. Rodrigues, J. C. S. Silva, P. L. S. de Oliveira, "CRISP-EDM: uma proposta de adaptação do Modelo CRISP-DM para mineração de dados educacionais," in Simpósio Brasileiro de Informática na Educação (SBIE), 1092–1101, SBC, 2020, doi:10.5753/cbie.sbie.2020.1092.

[42] K. Anguraj, B. Thiyaneswaran, G. Megashree, J. Preetha Shri, S. Navya, J. Jayanthi, "Crop recommendation on analyzing soil using machine learning," Turkish Journal of Computer and Mathematics Education (TURCOMAT), **12**(6), 1784–1791, 2021.

[43] J. An, "Using CatBoost and Other Supervised Machine Learning Algorithms to Predict Alzheimer's Disease," in 2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA), 1732–1739, 2022, doi:10.1109/ICMLA55696.2022.00265.

[44] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, **12**, 2825–2830, 2011, doi:10.5555/1953048.2078195.

[45] C. Liu, M. White, G. Newell, "Measuring the Accuracy of Species Distribution Models: A Review," in Proceedings of the 18th World IMACS / MODSIM Congress, 4241–4246, Victoria, Australia, 2009, reviewed accuracy metrics like discrimination capacity and reliability in SDMs; introduced both threshold-dependent and threshold-independent indices.

[46] M. Sokolova, G. Lapalme, "Learning Opinions in User-Generated Web Content," Natural Language Engineering, **17**(4), 541–567, 2011, doi:10.1017/S135132491100012X, first published online 11 March 2011; Cambridge University Press.

[47] M. Bekkar, H. K. Djemaa, T. A. Alitouche, "Evaluation Measures for Models Assessment over Imbalanced Data Sets," Journal of Information Engineering and Applications, **3**(10), 27–38, 2013, open access; ISSN 2224-5782 (print), 2225-0506 (online).

[48] R. L. Chambers, H. Chandra, N. Tzavidis, "On Bias-Robust Mean Squared Error Estimation for Pseudo-Linear Small Area Estimators," Survey Methodology, **37**(2), 153–170, 2011, develops a simpler, bias-robust MSE estimator for pseudo-linear small area estimators, including EBLUP, MBDE, and M-quantile predictors.

[49] T. Hastie, R. Tibshirani, J. Friedman, "The elements of statistical learning. Springer series in statistics," New York, NY, USA, 2001.

[50] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," Advances in neural information processing systems, **30**, 2017.

Advances in Science, Technology and Engineering Systems Journal Vol. 10, No. 4, 14-19(2025)

www.astesj.com

ASTESJ

ISSN: 2415-6698

The First Study on Ionospheric Peak Variability over Equatorial Africa (COSMIC-2)

Chukwuemeka Alexander. Osueke*, Obiageli Josphine Ugonabo, Tafon Williams Sivla, Akor John Yakubu, Eucheria Chidinma Okoro

Department of Physics & Astronomy, University of Nigeria, P.O. Box 3238, Nsukka 410001 Nigeria

(emekaosueke92@gmail.com,obiageli.ugonabo@unn.edu.ng,william.sivla@unn.edu.ng,john.yakubu@unn.edu.ng, eucharia.okoro@unn.edu.ng).

A R T I C L E I N F O

Article history:
Received: 28 May, 2025
Revised: 9 July, 2025
Accepted: 11 July, 2025
Online: 22 August, 2025

Keywords:
Ionosphere
COSMIC-2
NmF2
International Reference
Ionosphere Model
HmF2
Electron Density

A B S T R A C T

In regions like the African equatorial region, where ground-based sensors like ionosondes and incoherent scatter radars are limited, satellite-based radio occultation (RO) observations offer a new alternative for ionospheric data collecting and optimization. Using RO measurements from the mostly newly launched COSMIC-2 (Constellation Observing System for Meteorology, Ionosphere, and Climate-2) mission, hence, the equatorial Africa, ionospheric peak parameters were first studied and published. Data obtained over Ilorin (8.50ºN, 4.50ºE) and Abuja (8.99ºN, 7.38ºE) from October 2019 to December 2020 were utilized in this study. The results show that diurnal profiles of NmF2 and hmF2 have crests at around 10:00 to 18:00 UT and troughs at around 04:00 to 05:00 UT. The NmF2 values are usually less than $2×10^5$ cm^{-3} during the early morning hours before sunrise, and they can reach~ 6-10 $×10^5$ cm^{-3} after sunrise, depending on season. NmF2 and hmF2 exhibit peaks between 18:00 UT and 10:00 UT and troughs between 05:00 UT and 04:00 UT. For the seasonal variation, the NmF2 values can reach approximately 6–10 $×10^5$ cm^{-3} after sunrise, while they are often less than $2×10^5$ cm^{-3} in the early morning hours before sunrise. We recorded higher values in the equinoxes (March and September) season than in the solstices (June and December) season. The 2016-IRI model overestimates the COSMIC-2 NmF2 measurements, typically during the decline phases of the 24-hour trend at 60% differences. During the March equinox season, the percentage RMSD (root-mean-squared deviation) is (~ 29%) smaller than the solstice seasons with (~30%). The percentage RMSD (root-mean-squared deviation) is the smallest. The disparity was so much observed during the September equinox, at a time when RMSD reaches ~ $1.7×10^5$ electrons/cm3 and the percentage RMSD reaches ~ 35%.

1. Introduction

The ionosphere is a region with the highest ionization in the Earth's atmosphere. It is identified as the upper part of the Earth's atmosphere in which free electrons and ions exist under the control of the gravity and magnetic field of the planet, and whose concentration is sufficient to affect the propagation of electromagnetic waves [1] and [2]. This region of the Earth's atmosphere extends from roughly 60 kilometers to 1000 kilometers in height [3]. It includes the mesosphere, thermosphere, and a portion of the exosphere [4]. Due to its electrification attribute, the ionosphere interacts and influences radio propagation through it [5]. Because the upper atmosphere and upper ionosphere are co-located, the very high temperatures in the upper atmosphere persist in the upper ionosphere, which is a result of X-ray radiation from the Sun [6]. According to some researchers [7], [8], [9], [10] and [11]. Ionosphere is said to be critical and essential in telecommunication because it allows radio waves to travel both short and long distances on Earth. It also affects has been studied extensively using different satellite-based and ground-based technologies. Ionosondes/digisondes are one of the most reliable sources of ionospheric electron density (edensity) information. However, in the African equatorial region due to relatively high costs of installation and maintenance the mentioned instrument for measuring ionosphere is unavailable. Satellite-based systems, like the COSMIC (Constellation Observing System for Meteorology, Ionosphere, and Climate) mission, offer good opportunities to acquire ionospheric density information over the African equatorial region through radio occultation (RO) technique. The technique has the advantages of global coverage, high vertical resolution, and the absence of signal disturbances caused by the troposphere [12]. When a low-Earth orbit receiver (like the

*Corresponding Author: Osueke, Chukwuemeka Alexander Department of Physics & Astronomy, University of Nigeria, P.O. Box 3238, Nsukka 410001 Nigeria, +2348135752624, emekaosueke92@gmail.com

https://dx.doi.org/10.25046/aj100402

COSMIC satellite) observes a rising or setting GNSS satellite, a GNSS-RO takes place [13]. Because atmospheric refractivity modifies the occulted GNSS signals, this signal is examined to provide data on ionospheric electron density and other atmospheric parameters like density, pressure, temperature and humidity [14], [15] and [16].

The RO technique was first demonstrated with the GPS/MET experiment [17] and [18] and its success has led to applying the technique in many more satellite missions, including the COSMIC mission. Several authors have used the COSMIC data such as [19], [20], [21] and [22] to study the variation of ionospheric electron distribution in various parts of the globe. The peak value of edensity, NmF2, and the corresponding height of peak edensity, hmF2, do characterize electron density distribution in the ionosphere, and they are therefore key parameters for modeling the edensity distribution in the ionosphere. The COSMIC satellites are low-earth orbiting satellites with an altitudinal resolution. It varies depending on the specific implementation and instrumentation, but it is typically on the order of 0.5-1.4 kilometers (0.3-0.9 miles) [23]. These studies were based on the total electron content (TEC) parameter integrated from the COSMIC-2 measurements. There is a particularly lacking research report on the behavior of RO-derived density measurements in the African equatorial region, and much more, using the newly launched COSMIC-2 mission measurements. These kinds of studies are necessary to provide scientists/researchers with vital information on the performance of such measurements. The purpose of this research is to use RO density measurements obtained from the COSMIC-2 mission to characterize the ionospheric edensity in the equatorial African region. This is the first study employing RO electron density measurements from the COSMIC-2 mission to characterize the African equatorial ionosphere.

2. Data and Methods

Data was obtained from the University Corporation for Atmospheric Research (https://data.cosmic.ucar.edu/gnss-ro/) in the form of ionPrf files (second-level processed files). Data is available on this website starting from the 1st of October 2019. All available data from the 1st of October 2019 to the 7th of December 2020 were used. Data obtained from two stations in Nigeria, specifically Abuja (8.99ºN, 7.38ºE) and Ilorin (8.50ºN, 4.50ºE) was used in this study. We considered only observations falling within a circular window of radius 5 great circle degrees around each of the stations. Ilorin ionosonde measurements were used for comparison with corresponding COSMIC measurements. The hourly values measurement for diurnal variabilities of NmF2 and hmF2 over the seasons for each season was calculated. The seasons were constituted as follows: March equinox (March and April), June solstice (May, June, July and August), September equinox (September and October), and December solstice (January, February, November and December) according to [24]. DST data was obtained from OmniWeb (https://omniweb.gsfc.nasa.gov/form/dx1.html) only data for geomagnetic quiet days (|DST| < 30 nT) was used for this computation.

The median absolute deviation (MAD) processing of the NmF2 and hmF2 variabilities expressed in equation (1) was computed as a measure of the typical deviations from the median values.

$$MAD = median(|X_i - \hat{X}|) \tag{1}$$

Where X_i are elements in the distribution, and $\tilde{X}$ is median of the distribution

$$RMSD = \left(\frac{\sum_{i=1}^{n}(C_i - I_i)^2}{n}\right)^{0.5} \tag{2}$$

Where C_i and I_i are respectively corresponding COSMIC and IRI model values, and n is the number of samples.

3. Results and Discussion

The results presented in this paper are in two scenarios; in the first scenario, the altitudinal variation of electron density is presented, and in the second scenario, the diurnal and seasonal variations of the NmF2 and hmF2 parameters are presented.

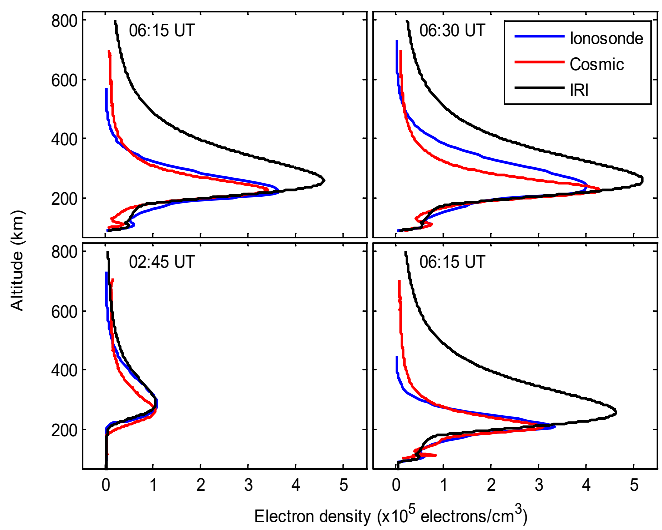

Figure 1: Altitudinal Variation of Ionospheric Electron Density

Figure1. Electron density (edensity) profiles obtained from COSMIC-2 mission, Ionosonde and 2016-IRI model for Ilorin. The edensity profiles on the top panel are respectively for 06:15 UT and 06:30 UT on day number 274 (October 1), year 2019, while the electron density profiles on the bottom panel are respectively for 02:45 UT and 06:15 UT on day number 276 (October 3), year 2019.

Figure 1 shows the altitudinal variations of edensities at four different times in two days when there was a record of coincident edensity measurements from both COSMIC and ionosonde equipment. The edensity profiles on the top panel are respectively for 06:15 UT and 06:30 UT on day number 274 (October 1), the year 2019, while the edensity profiles on the bottom panel are respectively for 02:45 UT and 06:15 UT on day number 276 (October 3), the year 2019. Local time in Nigeria is UT+1 hour. It is important to clarify here that we would have been interested to show these results from different times of the day, but we are limited by the number of times when there are coincident observations from both COSMIC and ionosonde equipment.

Throughout the study investigated in this research, there was available data from the ionosonde only during the few days between days numbers 274 and 277. During these few days, coincident COSMIC measurements were recorded during the times illustrated. We considered only COSMIC measurements recorded within a spatial window of 5 great circle degrees around the Ilorin ionosonde station and within a time window of 15 minutes of the ionosonde measurement.

Figure 1 illustrates the typical pattern of electron density variation with altitude at the location. The figure shows that the edensities typically peak at about 250 – 300 km. It is evident from the figure that the IRI model overestimates the peak parameters (both the peak edensity value and the height of peak edensity), especially for the profiles around 06:15 and 06:30 UT. Panels 2 and 3 in Figure 1 provide more premises for analyzing the IRI model predictions of the peak parameters. The patterns provided by the COSMIC and ionosonde measurements are more identical, except for a case like the 02:45 UT measurement where the ionosonde measurements are slightly greater than the COSMIC measurements at around 300 – 400 km altitudes. These slight differences could be attributed to the reason that the ionosonde measurements above the F2 peak altitude are modeled values, rather than ‘pure’ measurements. The inherent modeling errors could be responsible for the slight departures noted at altitudes above the F2 peak. The good level of consistency between the ionosonde and COSMIC measurements indicates that the measurements can both be independently used to study ionospheric electron density variations in the region. Given the paucity of ionosonde data for the period of this study, we next present an extended climatologic study of the peak parameters that are based on COSMIC-2 measurements. This ionospheric variability study is the first COSMIC-2-based measurement of ionospheric peak parameters carried out in the region. Finally, the few discrepancies observed in our data between the Cosmic-2 measurement and IRI-model above the F2 peak, where the ionosonde measurement slightly overestimates Cosmic-2 densities, are attributed to the modeled nature of ionosonde outputs at those altitudes. This insight emphasizes the importance of cautious interpretation of ionsonder data above the F2 peak and supports the use of satellite-based data for capturing high altitude ionospheric structure. There was an indication that the IRI model overstatement during the early hours of (06:15 UT and 06: 30 UT). This discrepancy highlights a limitation in the IRI model's ability to accurately capture the local ionospheric condition in equatorial West Africa. Therefore, we suggest that regional calibration of the model may be necessary to improve its performance in the zone.

Figure 2 shows diurnal variation patterns of the NmF2 at Abuja for the different seasons. For each season, hourly based median values and MADs of the NmF2 were computed as described in the Data and Methods section. In Figure 2 the NmF2 daily patterns show troughs at about 04:00 UT, just before sunrise. The figure also shows that the NmF2 values are frequently less than 2×105 cm^{-3} during the early morning hours before sunrise. The peaks of the diurnal NmF2 profiles are about 6-10 ×105 cm^{-3}, and these typically occur at around 16:00 UT. The day-time values (after sunrise and before sunset) are usually greater than 4.5×105 cm^{-3}, and the night-time values usually drop below 2×105 cm^{-3}. Seasonally, equinoxes (March and September) have higher NmF2 values than during the solstices (June and December). This is because the location is in the equatorial region; the equatorial region receives more direct sunlight during the equinoxes than during the solstices. It is also evident from Figure 2 that there is good agreement between the NmF2 profiles produced by the IRI model and COSMIC-2. The patterns of the profiles are similar. Table 1 contains statistics on the relationship between the IRI model and COSMIC-2 profiles. The table shows that correlation coefficients between the profiles are greater than 0.9 for each of the four seasons. The values reveal that there is the greatest similarity in the patterns of the profiles during the June solstice. The IRI and COSMIC-2 NmF2 values are frequently observed to be within the error limits of each other. However, there are occasional disparities between the two datasets. Table 1 contains information on the root-mean-squared differences (RMSDs) and associated percentage differences between the two datasets. The RMSDs, computed using the formula in equation (2), are measures of the differences between the two datasets; they indicate the typical differences between the datasets in units of electron density, while the percentage difference indicates the difference as a percentage of the electron density at the given instances.

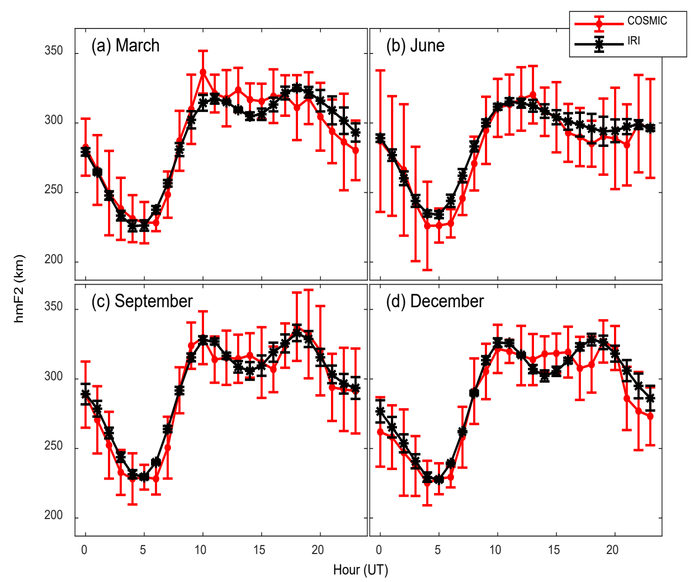

Figure 2: Diurnal Profiles of the NmF2 from COSMIC-2 measurements and IRI model for (a) March Equinox, (b) June Solstice, (c) September equinox, and (d) December Solstice, during the period of data used in this study.

Smaller values of the RMSD and the percentage difference therefore connote better agreement between the two datasets. The table, as well as Figure 2, show that smaller RMSDs (root-mean-squared differences) are obtained during the solstice seasons (~ 1.1×105 electrons/cm3). These are therefore the seasons when the IRI model values have better agreement with the COSMIC-2 measurements. The smallest RMSD is particularly recorded during the June solstice season. Considering that it is also in the June season that the two datasets have the greatest correlation, we note that the IRI model most accurately reproduces COSMIC-2 measurements during the June solstice.

Table 1: Table of correlation coefficients, RMSDs, and percentage RMSDs, computed between COSMIC-2 NmF2 measurements and IRI model values

Season	Correlation coefficient	RMSD ($\times 10^5$ electrons/cm^3)	Percentage RMSD (%)
March equinox	0.96	1.23	29.01
June solstice	0.99	1.07	31.64
September equinox	0.95	1.67	35.78
December solstice	0.92	1.09	30.49

One valid argument is that the RMSDs are smaller for the solstices because the edensity magnitudes are correspondingly smaller for the solstice seasons [25]. This is why it becomes meaningful to consider the percentage difference, which is computed relative to the electron density magnitudes. Table 1 reveals that the percentage difference is rather smallest for the March equinox season (~ 29%) this reveals that the model's relative accuracy is seasoned dependent, not just dependent on the magnitude of electron density. This quantitative seasonal evaluation provides a finer resolution of model performance. The value is however closely followed by values for the solstices. The greatest disparity between the datasets is seen during the September equinox; it is during this season that we record the greatest values of RMSD (~ 1.7×105 electrons/cm^3) and percentage difference (~ 36%) between the COSMIC-2 and IRI datasets.

Figure 2 reveals that the IRI model typically overestimates the COSMIC-2 measurements. A similar scenario is noted in Figure 1, where the IRI model is observed to overestimate both ionosonde and COSMIC-2 electron density measurements. In Table 2, we provide indications of the RMSDs for different phases of a diurnal profile. We split a diurnal profile into three phases; a rising phase (04:00 to 08:00 UT), a peak phase (09:00 to 16:00 UT), and a decline phase (17:00 to 03:00 UT).

Table 2: Table of RMSDs and percentage RMSDs, computed between COSMIC-2 NmF2 measurements and IRI model values for different phases of a diurnal profile. The electron densities (in $\times 10^5$ electrons/cm^3) are outside the brackets while the percentage RMSDs (in %) are inside the brackets.

Season	Rising Phase	Peak Phase	Decline Phase
March equinox	0.49 (14.82)	1.19 (19.45)	1.47 (44.50)
June solstice	0.83 (27.20)	1.25 (23.98)	1.02 (46.70)
September equinox	1.10 (31.55)	1.38 (19.60)	2.04 (58.41)
December solstice	056 (19.95)	0.49 (08.95)	1.49 (61.84)

Table 2 shows that the RMSDs are usually lowest during the rising phase of the profile, except for the December solstice when the RMSD is lowest during the peak phase. On the other hand, the RMSDs are usually highest during the decline phase of the profile, except for the June Solstice when the RMSD is highest during the peak phase. During the rising phase, the RMSDs are in the range of ~ (0.5 – 1.1) $\times 10^5$ electrons/cm^3. The upper limit of the range is slightly greater during the peak phase (~ 1.4×105 electrons/cm^3), and the greatest RMSDs are observed during the decline phase; ~ (1.0 – 2.0) $\times 10^5$ electrons/cm^3. The percentage differences are however lower during the peak phase than during the rising phase, except for the March equinox. The percentage differences remain consistently greatest during the decline phase for all four seasons, indicating that this is the phase in which the IRI model values are mostly different from the COSMIC-2 measurements. A cursory look at Figure 2 reveals that it is during the decline phase that the IRI model conspicuously overestimates the COSMIC-2 measurements. Finally, there is a great difference occurring consistently during the decline phase across all the seasons, with percentage RMSDs reaching up to ~62%. This pattern shows a systematic overestimation by the IRI model during post-sunset hours, suggesting that the mode does not adequately capture the ionospheric decay phase at night, which is crucial for applications involving night-hour HF communication and satellite signal delay corrections.

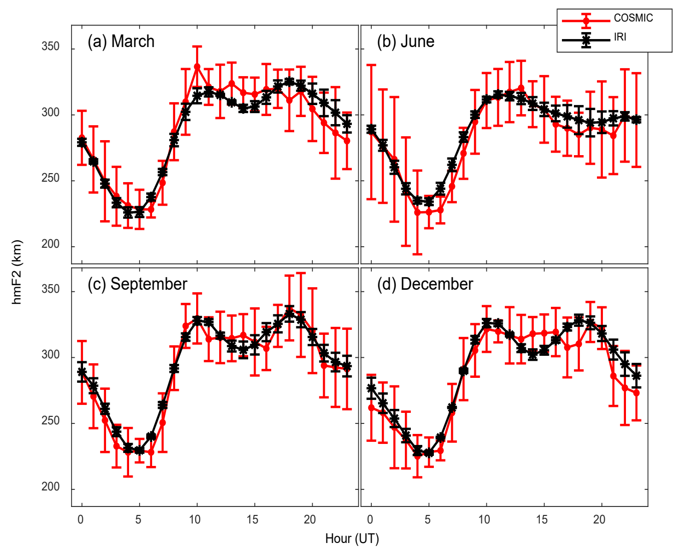

Figure 3: Diurnal Profiles of the hmF2 from COSMIC-2 measurements and IRI model for (a) March Equinox, (b) June Solstice, (c) September equinox, and (d) December Solstice, during the period of data used in this study.

Corresponding comparisons between COSMIC-2 hmF2 measurements and IRI hmF2 values are illustrated in Figure 3. The figure shows that the electron densities peak at altitudes between 200 and 350 km, depending on the time of the day and season. In a pattern that is somewhat similar to the NmF2, the altitudes of peak electron density are higher during the day (after sunrise) and lower at night (after sunset). The troughs of the hmF2 diurnal profiles usually occur at around 04:00 to 05:00 UT, while the crests occur between 10:00 and 18:00 UT. Figure 3 also shows that there is good agreement between hmF2 values from both systems; the IRI model values are well within the error limits of the COSMIC-2 measurements and the patterns of both profiles are identical, showing that the IRI model does effectively reproduce the COSMIC-2 hmF2 measurements. The correlation coefficients between the COSMIC-2 and IRI hmF2 values (shown in Table 3) are consistently greater than 0.95 for all seasons. The RMSD values are typically ~ 10 km and less, translating to percentage

RMSDs that are less than 4% for all four seasons. The results generally show that there is better agreement between the hmF2 values than between the NmF2 values.

Table 3: Table of correlation coefficients, RMSDs, and percentage RMSDs, computed between COSMIC-2 hmF2 measurements and IRI model values.

Season	**Correlation coefficient**	**RMSD (km)**	**Percentage RMSD (%)**
March equinox	0. 96	19.74	3.37
June solstice	0.98	7.79	2.79
September equinox	0.98	7.42	2.54
December solstice	0.96	10.33	3.61

4. Conclusion

First climatology of the NmF2 and hmF2 parameters using COSMIC-2 RO measurements in the equatorial African region is studied. There was the presentation of altitudinal electron density profiles from the COSMIC-2 mission, alongside corresponding ionosonde measurements and IRI model values. The results showed that the COSMIC-2 and ionosonde measurements were more closely related, while the IRI typically overestimated both measurements, especially at the F2 peak altitude and above. However, this study not only validates the use of COSMIC-2 data for ionospheric research in equatorial Africa but also brings attention to the need for model improvements and offers new perspectives on the vertical structure of the ionosphere in the region. These contributions are particularly valuable given the limited observational infrastructure available in West Africa equatorial and reinforce the potential of satellite-based measurements in supporting ionospheric monitoring and modeling efforts.

Diurnal profiles of the NmF2 and hmF2 were constructed from both COSMIC-2 and IRI on a seasonal basis using the medians of the values for each of the seasons, binned hourly.

The results generally showed a significant correlation (correlation coefficients greater than 0.9) between the COSMIC-2 and IRI model values. Typical differences between the COSMIC-2 and IRI NmF2 values were in the range of ~ 1.1 ×105 to 1.7 ×105 electrons/cm3, corresponding to percentage RMSDs of ~ 29% to 36%. The RMSDs were lower during the solstices and greater during the equinoxes, implying that the IRI model values were closer to the COSMIC-2 measurements during the solstices than during the equinoxes. This could be a result of the relatively lower electron densities recorded in the equatorial region during the solstices than during the equinoxes, and so the justification for computing the percentage RMSDs. The percentage RMSD was lowest for the March equinox season (~29%), and closely followed by values for the solstices (~30%). The greatest differences (both RMSD and percentage RMSD) were recorded during the September equinox (~35%). There was also an investigation of the differences between the COSMIC-2 measurements and the IRI values for three different phases of a diurnal profile (rising, peak, and decline phases). The results showed that the least differences (both RMSDs and percentage RMSDs) were usually obtained during the rising phase, while the greatest differences were usually obtained during the decline phase.

It was conspicuously observed that during the decline phase; the IRI model generally overestimated the COSMIC-2 NmF2 observations. The was a generally good agreement between the COSMIC-2 hmF2 measurements and the IRI model values, with RMSDs less than ~ 10 km and less, and percentage RMSDs less than 4%. Aside from the new results of electron density, NmF2, and hmF2 variations presented for the African equatorial African region using COSMIC-2 measurements, the paper also provides vital information on the climatologic differences between the IRI model predictions and the COSMIC-2 measurements, for improving ionospheric modeling in the region. However, the IRI model shows better in producing peak height (hmF2) than peak density (NmF2), this result will enhance understanding of model reliability across different ionospheric parameters.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Radio occultation electron density data from the COSMIC-2 mission was used in this work. The authors are grateful to the sponsors and operators of the COSMIC-2 mission; the National Science Council and National Space Organization (NSPO), the National Oceanic and Atmospheric Administration (NOAA), the University Corporation for Atmospheric Research (UCAR), the National Science Foundation (NSF), National Aeronautics and Space Administration (NASA). Ionosonde electron density data was obtained from DIDBASE GIRO (https://ulcar.uml.edu/DIDBase/) through the SAO Explorer software. The authors are grateful to the GIRO team for providing and granting access to the ionosonde data. The authors acknowledge the OMNIWEB (https://omniweb.gsfc.nasa.gov/form/dx1.html) for providing and granting access to DST data used in this work. The first author is grateful to the University Support Program (USP) of the Centre for Atmospheric Research (CAR) for the support they provided using R&D funds from the Federal Government of Nigeria, and to the University of Nigeria - Nsukka (UNN) for approving the request to visit the Space Environment Research Facility of CAR.

References

[1] N.P. Chapagain, L. Patangate, “Ionosphere and its influence in communication systems,” An annual publication of Central Department of Physics, 10, 2016.

[2] V.L. Bychkov, G.V. Golubkov, A.I. Nikitin, The atmosphere and ionosphere, Springer, 2010.

[3] P.K. Bhattacharjee, “Fundamental to electromagnetic waves,” International Journal of Trend in Scientific Research and Development, 7, 454-462, 2023.

[4] R.K. Cole, E.T. Pierce, “Electrification in the earth's atmosphere for altitudes between 0 and 100 kilometers,” Journal of Geophysical Research, **70**(12), 2735-2749, 1965.

[5] R.A. Vincent, “The dynamics of the mesosphere and lower thermosphere: a brief review,” Progress in Earth and Planetary Science, 2, 1-13, 2015.

[6] M. Atiq, "Historical review of ionosphere in perspective of sources of ionization and radio waves propagation," Research & Reviews: Journal of Space Science & Technology, **7**(2), 28-39, 2018.

[7] O.A. Ogunmodimu, Auroral Radio Absorption: Modelling and Prediction, Ph.D Thesis, Lancaster University (United Kingdom), 2016.

[8] A. Daniel, G. Tilahun, A. Teshager, "Effect of ionosphere on radio wave propagation," International Journal of Research, **3**(9), 65-74, 2016.

[9] E.V. Appleton, "Wireless studies of the ionosphere," Institution of Electrical Engineers- Proceedings of the Wireless Section, **7**(21), 257-265, 1932.

[10] K.G. Budden, The propagation of radio waves: the theory of radio waves of low power in the ionosphere and magnetosphere, Cambridge University Press, 1988.

[11] H. Sizun, P. de-Fornel, Radio wave propagation for telecommunication applications, 35-67, Springer, Berlin, 2005.

[12] S. Dubey, R. Wahi, A.K. Gwal, "Ionospheric effects on GPS positioning," Advances in Space Research, **38**(11), 2478-2484, 2006.

[13] D. Atlas, R.C. Beal, R.A. Brown, D.P. Mey, R.K. Moore, C.G. Rapley, C.T. Swift, "Problems and future directions in remote sensing of the oceans and troposphere: a workshop report," Journal of Geophysical Research: Oceans, **91**(C2), 2525-2548, 1986.

[14] A.J. Mannucci, C.O. Ao, W. Williamson, "GNSS radio occultation," Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications, 1, 971-1013, 2020.

[15] R. Notarpietro, M. Cucca, S. Bonafoni, "GNSS signals: a powerful source for atmosphere and Earth's surface monitoring," Remote Sensing of Planet Earth, 171-200, 2012.

[16] S. Jin, R. Jin, X. Liu, GNSS atmospheric seismology, Springer, Berlin/Heidelberg, Germany, 2019.

[17] R. Padullés, E. Cardellach, K.N. Wang, C.O. Ao, F.J. Turk, M.D.L. Torre-Juárez, "Assessment of global navigation satellite system (GNSS) radio occultation refractivity under heavy precipitation," Atmospheric Chemistry and Physics, 18(16), 11697-11708, 2018.

[18] Y.A. Liou, A.G. Pavelyev, J. Wickert, T. Schmidt, A.A. Pavelyev, "Analysis of atmospheric and ionospheric structures using the GPS/MET and CHAMP radio occultation database: a methodological review," GPS Solutions, 9, 122-134, 2005.

[19] C.O. Ao, G.A. Hajj, T.K. Meehan, D. Dong, B.A. Iijima, A.J. Mannucci, E.R. Kursinski, "Rising and setting GPS occultations by use of open-loop tracking," Journal of Geophysical Research, 114, D04101, 2009, doi:10.1029/2008JD010483.

[20] Y.H. Chu, C.L. Su, H.T. Ko, "A global survey of COSMIC ionospheric peak electron density and its height: a comparison with ground-based ionosonde measurements," Advances in Space Research, 45, 431-439, 2010.

[21] M.M. Hoque, N. Jakowski, "A new global model for the ionospheric F2 peak height for radio wave propagation," Annales Geophysicae, 30, 797–809, 2012.

[22] L. Hu, B. Ning, L. Liu, B. Zhao, Y. Chen, G. Li, "Comparison between ionospheric peak parameters retrieved from COSMIC measurement and ionosonde observation over Sanya," Advances in Space Research, 54, 929-938, 2014.

[23] E.R. Kursinski, G.A. Hajj, J.T. Schofield, R.P. Linfield, K.R. Hardy, "Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System," Journal of Geophysical Research: Atmospheres, **102**(D19), 23429-23465, 1997.

[24] S. Mukherjee, S. Sarkar, P.K. Purohit, A.K. Gwal, "Seasonal variation of total electron content at crest of equatorial anomaly station during low solar activity conditions," Advances in Space Research, **46**(3), 291-295, 2010.

[25] D. Okoh, H. JohnBosco, R. Babatunde, S. Gopi, B.W. Joshua, O. Joseph, O. Olivier, M.M. Tshimangadzo, "Storm-time modeling of the African regional ionospheric total electron content using artificial neural networks," Space Weather, **18**(9), e2020SW002525, 2020.

Advances in Science, Technology and Engineering Systems Journal Vol. 10, No. 4, 20-31 (2025)
www.astesj.com
ASTESJ
ISSN: 2415-6698

Complete System and Interactions of MMF Harmonics in a Squirrel Cage Induction Motor; Differential Leakage; Analytic Calculation

Gábor Kovács

Ganz Electric Works, 1095 Budapest, Hungary

A R T I C L E I N F O

Article history:
Received: 8 July, 2025
Revised: 10 July, 20205
Accepted: 4 August, 2025
Online: 25 August, 2025

Keywords:
Induction motor
Squirrel cage
MMF space harmonics
Differential leakage
Parasitic torques
Radial/tangential magnetic forces

A B S T R A C T

The importance of MMF space harmonics in squirrel-cage induction motors has been recognized in the literature since the beginning. Their details have been analyzed over the years, but only partly systematized. In this article, however, not only the origin and the entire system of that harmonics are described, but also their interaction causing the asynchronous parasitic torques, the synchronous parasitic torques, the radial and tangential magnetic force waves on a systematic way; all of this is done now targeting a complete system, so we include all harmonic phenomena in a single complete equivalent circuit diagram and a single table of the magnetic forces that occur. The so far missing equivalent elements causing the synchronous torques have now been defined. From these, the physical relationship between the phenomena that have often been examined separately will be clearly visible and can be easily calculated. Building on our previous studies, instead of the theoretically infinite number of harmonics hitherto considered, we will describe "exactly" the harmonics that really need to be taken into account, so that the number of harmonics will be much smaller than is usually considered by researchers and designers. Our formulae derived so far are successfully verified with the help of that completed equivalent circuit diagram, supplemented with the so far missing formula for tangential magnetic forces.

1. Introduction

The first complete description of MMF space harmonics occurring in an asynchronous machine is attributed in the literature to [1]. The first (and perhaps the only) series of measurements demonstrating the influence of MMF harmonics on the torque-speed characteristic curve, including the occurrence of synchronous machine phenomena in asynchronous machines, and which series of measurements has been referred to ever since, was carried out by [2]. The effect of MMF harmonics on noise was first systematized by [3]. The books in [4] and [5] not only provided a complete theoretical foundation for the asynchronous machine, becoming the number one fundamental works in the literature by not only defining those harmonics, but also by consistently tracing the effect and calculation of them. In [6], the author explicitly devoted their fundamental work to harmonics. In reality, however, they were, in our view, concerned always with a sub-area of the complete system. This is also true for [7], although he was the first to report the completeness of the MMF and also other harmonics occurring in the air gap of asynchronous machines, partly their interaction, further the sequence and complete procedure of the calculation; however, he did not make the number of harmonics to be taken into account dependent on the number of rotor slots, but thought of it as theoretically infinite.

Although [8] does not directly address their fundamental work to the topic of this article, it provides a theoretical approach to the basic laws of electrical engineering in relation to asynchronous machine, which is indispensable for our topic.

The investigations so far were complete in a certain sense, in a sub-area, indeed, it can always be stated, however, that something was always missing from the completeness, in our opinion the following three things.

The effect of the number of rotor slots was not taken into account, or not so as we believe it should have been. The next reason, which really completes our investigations, is that although previous investigations have recognized that the phenomena occurring in the machine can be modeled with a series of "small" asynchronous motors and "small" synchronous motors representing the effects of the MMF space harmonics, shaft-connected to the "harmonic-free" main motor [6], *the small*

*Corresponding Author: Gabor Kovacs, Ganz Electric Works, 1095 Budapest, Hungary, gabor.kovacs.ganz@gmail.com

https://dx.doi.org/10.25046/aj100403

synchronous motors, however*, have never been defined.* Nor have been recognized the relationship between the small synchronous motors used as models and the radial and tangential magnetic harmonic forces. The final reason is that – in the absence of our formulae – it has not been possible to do more than to specify the frequency of the harmful phenomena (for checking for resonance), so the designer could not define just the slot numbers to be chosen, what he was really looking for but only their opposite, the set of obviously most dangerous slot numbers to be avoided. The latter could only be examined as a way of "avoiding" the so-called slot harmonics.

2. Analytical model

As a model, such a machine is assumed for which the basic formulae for determining the usual resulting space harmonics are valid [6]: infinite relative permeability, two-dimensional fields without considering boundary and end effects, the machine consist of two smooth coaxial cylinders made of magnetic material, the cylinders are separated by the air gap, the conductors of infinitely small cross-section are located in the air gap.

In such a model, the distribution of the excitation current (no-load magnetizing current) flowing in the winding in the air gap will have the same shape as the shape of the magnetic induction or magnetic field, since there will be a linear relationship between the two. Therefore, we often do not talk about excitation (MMF), but about the air gap field and its harmonics. Considering its shape, the MMF curve is always "stepped", so it is natural that it contains harmonics.

If necessary, in addition to the MMF harmonics, reference will also be made to the magnetic conduction fluctuation harmonics arising due to the stator slot opening, but such case will be indicated explicitly. Time harmonics and harmonics of other origin, such important phenomena as eccentricity and saturation, are not included in the investigation, so that our examination can really concentrate on the MMF space harmonics. The supply voltage is always considered sinusoidal, and the motor is started by connecting it directly to the network (without a frequency converter drive).

The well-known equivalent circuit of the asynchronous machine is valid only for the fundamental harmonic magnetic field, therefore we develop an extended, special equivalent circuit that also takes into account the harmonic fields occurring in the case of the model outlined above.

With the help of our investigation, we explore the generation of parasitic torques during run-up, the generation of radial and tangential forces that cause vibrations and noise during operation, with the aim of understanding of origins and results of different interactions.

Since the derivations as well as the conclusions drawn from the model are based on the fundamental laws of electrical engineering, they do not require validation.

3. MMF harmonics in squirrel cage induction machines

MMF harmonics are given in first step for a 2p=2 pole machine.

Stator MMF harmonics

$$\nu = 2g \cdot m + 1 \tag{1}$$

where m stator phase number

$g=0, \pm1, \pm2, \pm3 \ldots$

the stator harmonics are *independent* of the stator slot number.

Rotor MMF harmonics

$$\mu_a = e\frac{Z_2}{p} + \nu_a \tag{2}$$

where Z_2 rotor slot number

p pole pair number

$e=0, \pm1, \pm2, \pm3 \ldots$

ν_a the stator harmonic that created this rotor harmonic

($Z_2/p=m_2$ rotor phase number)

the rotor harmonics are *dependent* of the rotor slot number.

Rotor harmonics created by the fundamental harmonic of the stator:

$$\mu_a = e\frac{Z_2}{p} + 1 \tag{3}$$

This description is suitable for general investigation of the machine - in the case of $q_1 = Z_1/2pm$ = integer i.e. without fraction/subharmonics - including the investigation of parasitic torques - and is therefore usually used there.

In fact, harmonics for the entire perimeter occur in the machine:

on the stator

$$\nu' = \nu p = 2m \cdot g \cdot p + p \tag{4}$$

on the rotor

$$\mu_a' = \mu_a p = e\frac{Z_2}{p}p + \nu_a p = e \cdot Z_2 + \nu_a p \tag{5}$$

Rotor harmonics created by the fundamental harmonic of the stator:

$$\mu_a' = e \cdot Z_2 + p \tag{6}$$

This notation should be used, for example, for the study of radial and tangential magnetic forces, because this is the only way to interpret the r-order of these force waves.

In addition to MMF (winding) harmonics, in the case of open or semi-closed slots on the stator, magnetic *conductivity* fluctuation induction harmonics are also generated in the air gap, the order number of which is necessarily determined by the combination of the slot number and the pole pair number.

Stator

$$\nu_{slot}' = g_1 \cdot Z_1 + p \tag{7}$$

where $g_1=\pm1$, (±2, ±3 …),

Rotor

$$\mu_a' = e_1 \cdot Z_2 + p \tag{8}$$

where $e_1=\pm1$, (±2, ±3 …).

Going back to stator MMF harmonics (1) let us highlight those of which $g=g_1 \cdot q_1$, where $q_1=Z_1/2mp$ (slot number per phase per pole)

$$\nu = 2mg_1q_1 + 1 = g_1 \cdot 2m\frac{Z_1}{2mp} + 1 = g_1 \cdot Z_1 / p + 1 = \nu_{slot} \tag{9}$$

$$\nu'_{slot} = \nu_{slot} p = g_1 Z_1 + p \tag{9a}$$

Expressions (7) and (9a) are identical, which means that the stator winding harmonics of order $g=g_1 \cdot q_1$ produce an order ν' along the circumference of the air gap that is equal to the order of the magnetic conductivity fluctuation harmonics. These parts of the whole set of MMF winding harmonics are therefore called slot harmonics. They are also characterized by the fact that their winding factor is the same as the winding factor of the fundamental harmonic $\xi_{\nu slot}=\xi_1$. The term "first (winding) slot harmonic" refers to the order $g_1=1$.

The above implies that the magnetic *conduction* fluctuation harmonics of open or semi-closed slots are "added" to, but only to, the *winding* harmonics of order $2mq_1 \pm 1$ and to its multiples $g_1 \cdot 2mq_1 \pm 1$; the two effects cannot then be separated during the operation of the machine. No new harmonics have appeared because of the effect of open slots.

From a comparison of (6) and (8), it is clear that all *rotor* MMF harmonics generated by the stator fundamental harmonic $\nu_a=1$ are slot harmonics. The magnetic conduction fluctuation harmonics of the rotor are of minor importance because of the small slot openings.

In the following, the initial definition of ν and μ_a will remain; if we still turn to using ν' and μ_a', it will be clearly indicated.

4. The reactance resulting from the MMF harmonics; air-gap leakage - differential leakage

Since the stator MMF is generated from the winding sides placed in the slots, it does not give a pure sinusoidal MMF: the MMF will have spatial harmonics. These harmonic fields are caused by the same stator current, which causes the so-called main/fundamental flux; but the speed of travel of the harmonic fields is not the same as the speed of travel of the main field, but slower being inversely proportional to the order number of that. For this reason, however, they induce a voltage in the winding with a frequency exactly corresponding to the network frequency, i.e. they behave just like a leakage flux. Hence their name, although they are not, is still *leakage* reactance. And since the phenomenon originates from processes occurring in the air gap, they are called "air-gap" leakage (Luftspaltstreuung) in several languages. The term "differential leakage" just in English comes from the fact that the air-gap leakage fields resulting from the harmonics can be thought of as the difference between the total field and the fundamental field. Since the frequency of the voltage induced in this way is the same as the frequency of the supply voltage, they can be inserted into the usual equivalent circuit diagram in the same way as any other (real) leakage reactance.

The air-gap leakage reactance is defined as the sum of an infinite series:

$$X_{s\sigma_1} = \sigma_1 X_m \tag{10}$$

where

$$\sigma_1 = \sum_{\nu \neq 1}^{\infty} \frac{1}{\nu^2} \frac{\xi_\nu^2}{\xi_1^2}$$

This phenomenon also occurs on the rotor, the corresponding reactance can, therefore, be included in the equivalent circuit diagram of the rotor. The air gap leakage fluxes also move slower in inverse proportion to the order number relative to the rotor, so they induce the same frequency in the rotor as the main flux (fundamental harmonic flux) and as the other leakage reactances do, so it is also justified to include them in the equivalent circuit diagram. Definition:

$$X'_{s2\sigma_1} = \sigma_2 X_m \tag{11}$$

where

$$\sigma_2 = \sum \frac{1}{\mu^2}$$

Instead of the inconvenient infinite summation, a formula was developed based on the Görges diagram.

The differential leakage reactance (let us remain with this term from now on) [5] (258)

$$X_{2\sigma} = X_m \cdot \sigma_{2\sigma} \tag{12}$$

where

$$\sigma_{2\sigma} = \frac{1}{\eta_2^2} - 1 \ ,$$

where

$$\eta_{2,1} = \frac{\sin\frac{p\pi}{Z_2}}{\frac{p\pi}{Z_2}} .$$

$\eta_{2,1}$ is hardly different from 1: $\eta_{2,1} \approx 1$.

Each harmonic flux of the differential leakage of the stator induces voltage (and current) in the rotor in the same way as the main (fundamental harmonic) flux - in other words, the cage rotor responds to a rotating field of any number of poles, i.e. responds to all MMF harmonic fields - and can therefore be imagined and modelled with small asynchronous machines shaft-connected to the rotor of the harmonic-free main motor. For this reason, it is usual to unfold the differential leakage elements (10) and assign to them one by one a separate small asynchronous machines' equivalent circuit diagram, electrically in series connection, mechanically in shaft connection, see e.g. [6] p. 110 Figure 41, but anywhere in the literature. These can then be used to explain and calculate the asynchronous parasitic torques and harmonic attenuation.

It has been shown that the leakage reactance of these small harmonic induction machines includes only the differential leakage reactance (of *their* rotor); this finding is only approximately true for the 5th and 7th harmonics [1], [5], [6], but in order to simplify the physical picture and the calculation, we consider them generally true for the time being.

The rotor differential leakage calculation of the small harmonic asynchronous machines was also developed based on the Görges diagram instead of the infinite series summation.

The differential leakage reactance [5] (268)

$$X_{2\sigma\nu} = X_{m\nu} \cdot \sigma_{2\sigma\nu} \tag{13}$$

where

$$\sigma_{2\sigma\nu} = \frac{1}{\eta_{2\nu}^2} - 1$$

rotor differential leakage coefficient.

The denominator is the so-called Jordan's coupling factor. Its definition ([5] 268b):

$$\eta_{2\nu_a} = \frac{\sin(\nu \frac{p\pi}{Z_2})}{\nu \frac{p\pi}{Z_2}} \tag{13a}$$

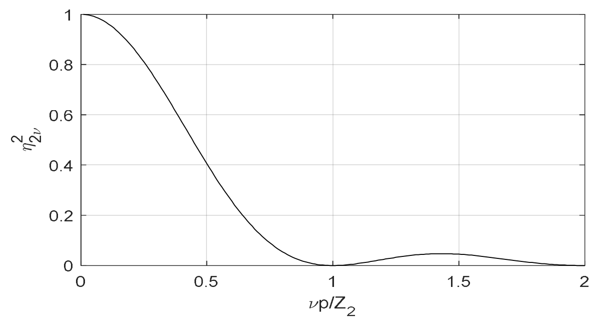

Figure 1: Plotting the value of $\eta_{2\nu}^2$ as a function of $\nu p/Z_2$ ([5] p.154. Figure 107.)

Expression (13a) is plotted on Figure 1. It shows how much the rotor responds to a stator harmonic. A zero or very low value indicates that the rotor does not respond to that harmonic. Then the differential leakage factor will be very high or even infinite ([6] Figure 17. p. 44).

Figure 1 is very difficult to interpret for a daily design work. Therefore, for a better illustration and more in line with the designer's thinking, we have already transformed this figure in [9]; the range $\nu p/Z_2>1$ was replaced by zero, since the highest value in this range is only ≈4.5%.

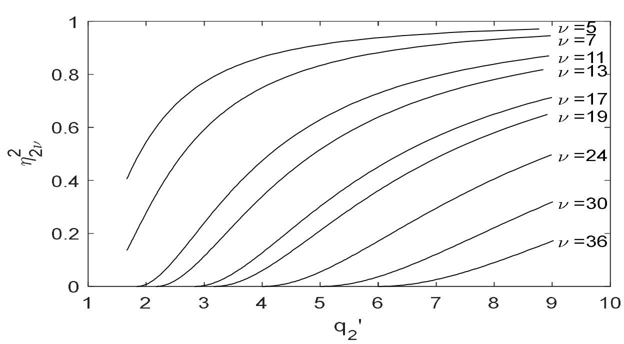

Figure 2: Representation of the value of $\eta_{2\nu}^2$ as a function of q_2' with ν as a parameter. Higher odd harmonics are replaced mathematically by adjacent (actually with 3-phase not existing) even harmonics only for better transparency

The "physical message" of the figure is the following:

if $q_2' \leq 2$, the machine does not respond to ν_a = 11, 13 (being otherwise the first slot harmonics of a stator winding q_1=2) and to any harmonics of higher order;

if $q_2' \leq 3$, the machine does not respond to ν_a = 17, 19 (being otherwise the first slot harmonics of a stator winding q_1=3) and to any harmonics of higher order; and so on.

This is the basis for our consideration of how to limit the theoretically infinite number of stator harmonics in the machine to a finite, not even unmanageably large, number of harmonics depending on the number of rotor slots, while still remaining theoretically correct. A first practical implementation of this approach has already been presented, dividing the stator differential leakage components into groups based on their identical properties, in [10] Figure 1. This figure contains an infinite series of differential leakage reactances, represented in the usual columnar manner; but unlike earlier representations no small asynchronous machine is assigned to stator harmonics higher than first slot harmonics. Now they are shown to the left of the current Figure 3, also in a columnar arrangement.

However, we now make a significant addition to the previous equivalent circuit diagrams (e.g. fill in the so far missing elements) by unfolding the differential leakage reactances of the *rotor* similarly to those of the stator, connected in series, also in a columnar manner, to the right of the fundamental equivalent circuit. The corresponding fluxes are also plotted with the voltage drop across the indicated reactance caused by the rotor fundamental harmonic current.

The fluxes represented by these reactances do not interact with the stator of the main harmonic-free motor. The normal stator does not respond to them, and this is because the normal stator is designed for a single fixed number of poles and would therefore only respond to a rotor current harmonic field of the same number of poles, i.e. μ=1 order. A cage rotor, however, *does not produce* such a wave; but if it did in rare cases, it would only produce it with a large ν_a order, [5] p. 153. Therefore, a small circuit like a small asynchronous machine, but pointing from the rotor to the stator, cannot be assigned to it, and cannot interact with it as an asynchronous machine would.

The flux fields represented by the differential leakage reactances of the rotor above, however, do interact with *each* of the MMF harmonic fields of the stator (and their associated current layer waves) and since they are independent of each other - i.e., neither is excited by the other - the interaction can only be synchronous in nature.

In accordance with the spirit of the figure, and more in accordance with the physical content, the quantities are proposed by the author to be referred with the term "harmonic" in a manner *equivalent* to the commonly used terms:

- harmonic asynchronous machine / synchronous machine
- harmonic asynchronous / synchronous torque
- harmonic radial/axial forces
- harmonic rotor current $I'_{2\nu}$

The usual term "parasitic" obviously refers to the harmful nature of the torque, the term "small" is not very expressive.

Another important phenomenon is that the spatial harmonics of the rotor's fundamental current (and also the spatial harmonics of the harmonic rotor currents see later) induce voltages in the stator, but their frequency is *not* the mains frequency, and therefore they do not appear in the usual voltage equations of

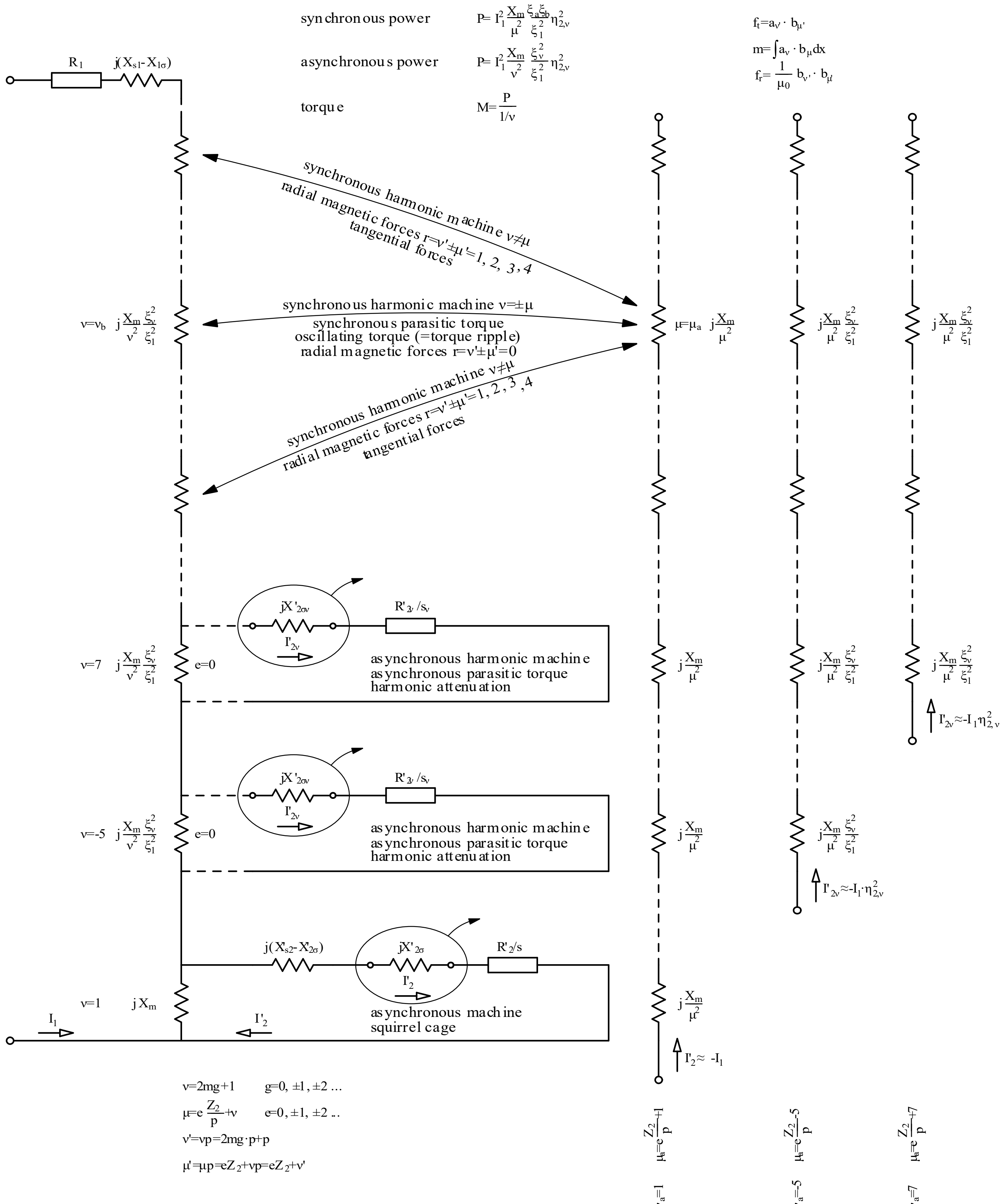

Figure 3: Single harmonic equivalent circuit of squirrel cage induction machine

the stator and hence not in the equivalent circuit diagram of an asynchronous machine. For the magnitude and frequency of the stator induced voltage, see Chapter 7 of [11]. Frequency of the induced voltage:

$$(1+\alpha)\cdot f_{net}$$

where $\alpha = eZ_2/p\cdot(1-s)$
where s slip of rotor
e positive integer

At a single slip value, at s=1 standstill, however, the frequency of the voltage induced in the stator is exactly the network frequency. Only in this case, the main field and the differential leakage fields cannot be distinguished [5].

Here we note that the cage rotor, however, *can generate* a wave of order $\mu=0$ for certain numbers of rotor slots with certain ν_a stator harmonics. In this case, the magnitude of the voltage drop $I_{2\nu}'\cdot X_m/\mu_a^2\cdot\xi_a^2/\xi_1^2$ in the affected "column" would be infinite, since $\mu=0$ is in the denominator; it is then the harmonic current $I_{2\nu}'$ of the small asynchronous motor affected by this phenomenon that eventually becomes zero. At the same time, $\eta_{2\nu}^2$ in the expression $I_{2\nu}'\approx -I_1\eta_{2\nu}^2$ will also be zero, so the two approaches give the same result. The physics of the phenomenon is that the wavelength of that stator harmonic is then exactly equal to the rotor's slot pitch, the induced voltages in the rotor bars are of the same magnitude and direction (zero sequence phenomenon), and thus cannot generate a current in the cage.

The newly applied consideration above (regarding the rotor differential leakage elements of the harmonic-free main machine) will be *further extended to the circuits of the small harmonic asynchronous machines*, which represent the effect of the stator's harmonic MMF on the cage: their leakage as differential leakage of *that* rotor can be unfolded, extracted and plotted separately in the same columnar manner alongside the column of the fundamental circuit. They also interact with the harmonic flux of *each* MMF in the stator, and form small harmonic synchronous machines in the same way.

Note that an asynchronous circuit ν and a rotor MMF harmonic μ may in some cases have the same order, and both may exist simultaneously. However, the physics of the two are different: one is a real (dependent) rotor *current* of order ν, the other is an independent harmonic rotor *flux field* of order μ, and therefore their behavior is different: the one is asynchronous, the other one is synchronous in nature.

All these "of synchronous" interactions of the fields can be perceived in the behavior of the machine as synchronous parasitic torque at well-defined rotor speed (including standstill) and as oscillating torque of the same small synchronous machine in the "out-of-synchronism" status over the rest of run-up speed range and at rated operation (acc. to the author incorrectly called torque ripple), and can also be perceived as radial and tangential magnetic forces over all the speed range (but deeply analyzed usually at no-load and rated load only).

The rotor harmonic fluxes close on the same magnetic paths as those of the stator in the case of the same order number; the calculation of their reactance is, therefore, *identical* to the calculation of the reactances representing the MMF harmonics of the stator see Figure. 3. With this statement, *the small harmonic synchronous motors are defined.*

The number of small synchronous machines to be included in the calculation depends on the followings:

- depending on the harmonic wavelength and the size of the (double) air gap, higher harmonics no longer reach the other side of the air gap, i.e. there is no interaction between the stator and rotor at these (high) harmonics; the highest order number of harmonics is "limited" in this way.
- the number of small asynchronous machines created is already limited as a function of the rotor slot number acc. to Figure 2.

Therefore, the number of harmonics to be considered is not infinite, even in theory.

For example, if the stator winding is chorded and the relative rotor slot number is $q_2'\leq 2$, then for two reasons, even based on the most rigorous theoretical considerations, it is sufficient to analyze the machine by considering the harmonics of the *fundamental* rotor current MMF only.

Figure 3 provides a complete map of the harmonic phenomena occurring in a short-circuited asynchronous machine. Every stator harmonic shown on the map is in harmonic relationship with every rotor harmonic. This means a very large number of interactions, although not an infinite number. The limiting phenomena are as follows:

- synchronous parasitic torque can only occur between harmonics of the same order: $\nu_b=\pm\mu_a$
- research in this direction has long proven that the order of the radial magnetic force waves to be considered is limited: $r=\nu_b'\pm\mu_a'\leq 4$ (for large machines ≤6) only
- the order of the magnetic force waves to be considered from the perspective of tangential forces has not been defined yet.

Taking this into account – and for the sake of clarity and transparency – Figure 3 shows only one harmonic pair interaction generating synchronous parasitic torque as an example, namely one in which the rotor harmonic happens to be the harmonic generated by the fundamental rotor current MMF (since such rotor harmonics *always* influence the behavior of the asynchronous machine), and the stator harmonic happens to be of such a high order that a small asynchronous machine no longer belongs to it and is therefore no longer attenuated. For the sake of clarity, we have only presented two interactions that generate radial and tangential magnetic force waves $r\neq 0$, which are expected to belong to the "neighboring" harmonics of the harmonic pair generating the synchronous parasitic torque.

As a result, when examining a short-circuited rotor induction motor, while maintaining theoretical rigor, we can limit ourselves to examining a few, but at least a manageable number of well-defined field harmonic interactions (see the studies in [11] Table 9).

Using Figure 3 as a map, *we do not need* to calculate the total, *resulting magnetic field* generated in the air gap. For the purposes of this topic, it is sufficient to consider only those harmonics that have a "counterpart on the other side."

5. Power, torque

The topic was investigated in [10] and [12]. Now, it is elaborated in a different way.

From Figure 3, the powers and torques that occur when the individual harmonics of the stator and rotor interact with each other can also be read.

5.1 Synchronous parasitic torque

From the equations of the synchronous machine the following is chosen to calculate the power of the small synchronous machine:

$$P = \frac{U_s \cdot U_r}{X_d} \tag{14}$$

where

$$U_s = I_1 X_s = I_1 \frac{X_m}{\nu_b^2} \frac{\xi_{\nu_b}^2}{\xi_1^2} \tag{14a}$$

the voltage corresponding to the stator harmonic MMF,

$$U_r = I_1 X_r = I_1 \eta_{2\nu_a}^2 \frac{X_m}{\mu_a^{\ 2}} \frac{\xi_{\nu_a}^2}{\xi_1^2} \tag{14b}$$

the voltage corresponding to the rotor harmonic MMF.

Regarding the harmonics of the rotor fundamental harmonic current;

$$U_r = I_1 X_r = I_1 \frac{X_m}{\mu_a^{\ 2}} \tag{15}$$

due to $\eta_{2,1}^2 \approx 1$ and per definition $\frac{\xi_{\nu_a}^2}{\xi_1^2} = 1$

Synchronous reactance X_d because good coupling is assumed between stator and rotor harmonic inductances:

$$X_d \approx \sqrt{X_s \cdot X_r} = \sqrt{\frac{X_m}{\nu_b^2} \frac{\xi_{\nu_b}^2}{\xi_1^2} \frac{X_m}{\mu_a^2} \frac{\xi_{\nu_a}^2}{\xi_1^2}} = \frac{X_m}{\mu_a^2} \frac{\xi_{\nu_b} \xi_{\nu_a}}{\xi_1^2} \tag{16}$$

where per definition: $\nu_b = \pm\mu_a$

Substituted

$$P = I_1^2 \eta_{2\nu_a}^2 \frac{\frac{X_m}{\nu_b^2} \frac{\xi_{\nu_b}^2}{\xi_1^2} \frac{X_m}{\mu_a^2} \frac{\xi_{\mu_a}^2}{\xi_1^2}}{\frac{X_m}{\mu_a^2} \frac{\xi_{\nu_b} \xi_{\nu_a}}{\xi_1^2}} = I_1^2 \eta_{2\nu_a}^2 \frac{X_m}{\mu_a^2} \frac{\xi_{\nu_b} \xi_{\nu_a}}{\xi_1^2} \tag{17}$$

During further transformations, the formula for the synchronous parasitic torque (=oscillating torque) that we derived earlier in a different way, will be obtained and confirmed [12].

5.2 Asynchronous parasitic torque

Power

$$P = U_s I_{2\nu}^{'} = I_1^2 \frac{X_m}{\nu^2} \frac{\xi_\nu^2}{\xi_1^2} \eta_{2\nu}^2 \tag{18}$$

During further transformations, the formula for the asynchronous parasitic torque that we derived earlier in a different way will be obtained and confirmed [10].

6. Noise measurements and torque oscillation measurements during start-up

It is a field experience of a commissioning engineer that while starting the motor at rated voltage (unlike starting at strongly reduced voltage in the testroom), significant noise occurs even at very low speeds, and then its frequency increases linearly, while the noise magnitude remains unchanged (apart from the air born noise); see “Tongeraden” [3]. After reaching the breakdown torque, the noise suddenly decreases significantly. It is not customary to measure it, but the frequency of the torque oscillation is obviously the same, and the magnitude of it (not losing sight of the logarithmic nature of the noise) behaves in the same manner as that of the noise. However, the analysis and especially the measurement of the latter is usually omitted from research.

Any measurements made during start-up would require a very significant apparatus, and their practical use is at the same time doubtful. However, there is no doubt that for the sake of a better understanding and especially for the proof of the theory of the processes, taking into account Table 1 see later, such measurement might still not be pointless.

7. The complete system of harmonic tangential and radial magnetic forces and torques occurring in the machine

For a better demonstration and understanding of the physical processes, the complete system of the forces created by the action of the MMF harmonics was included in Table 1 [13]. The table was prepared using the derivations of [8].

Let us see the physical background of the relationship between tangential forces (including synchronous torques) and radial forces. Both the flux density waves and the current layer waves of the stator and rotor resp., as a function of time and location (periphery) are written as follows:

$$\begin{aligned} b_s &= B_s \sin(\omega_s t - p_s x) \qquad & b_r &= B_r \sin(\omega_r t - p_r x - \varphi) \\ a_s &= A_s \cos(\omega_s t - p_s x) \qquad & a_r &= A_r \cos(\omega_r t - p_r x - \varphi) \end{aligned} \tag{19}$$

where ω_s and ω_r angular velocities of the *fields.*

There is the only stipulation that there is a causality relationship between the quantities within the stator as well as within the rotor:

$$a = const \frac{db}{dx} \tag{19a}$$

Otherwise, both stator and rotor quantities can be of any origin; between them there is no forced relationship.

When examining electrical machines in general, it is appropriate to look for the conditions for the formation of a constant torque in time. However, for the purpose of subject research, we are just looking for what happens when the conditions are not met. After all, all combinations occur during the interaction of stator and rotor harmonics.

Table 1: Complete system of radial and tangential harmonic magnetic forces (and torques) created in a squirrel cage induction machine

	Tangential force - Torque	Radial force
Basic formulae	$f_{tang} = B_s A_r \sin(\omega_s t - p_s x)\cdot\cos(\omega_r t - p_r x - \varphi) =$ $= 1/2\cdot B_s A_r[\sin((\omega_s-\omega_r)t-(p_s-p_r)x+\varphi)+$ $+\sin((\omega_s+\omega_r)t-(p_s+p_r)x-\varphi)]$	$f_{rad} \approx (B_s+B_r)^2 = B_s^2 + B_r^2 + 2B_sB_r =$ $= B_s^2/2\cdot(1-\cos(2\omega_s t - 2p_s x)) +$ $+ B_r^2/2\cdot(1-\cos(2\omega_r t - 2p_r x - 2\varphi)) +$ $+2B_sB_r * 1/2\cdot[(\cos((\omega_s-\omega_r)t-(p_s-p_r)x+\varphi)-$ $-\cos((\omega_s+\omega_r)t-(p_s+p_r)x-\varphi)]$
$p_s \neq p_r$ $\omega_s \neq \omega_r$	$f_{tang} = 1/2\cdot B_s A_r[\sin((\omega_s-\omega_r)t-(p_s-p_r)x+\varphi)+$ $+\sin((\omega_s+\omega_r)t-(p_s+p_r)x-\varphi)]$	$f_{rad} \approx ... + B_sB_r[\cos((\omega_s-\omega_r)t-(p_s-p_r)x+\varphi)-$ $-\cos((\omega_s+\omega_r)t-(p_s+p_r)x-\varphi)]$
	f_{tang} tangential force waves along the periphery as function of the position on the periphery and the time $M = -lr^2\int_0^{2\pi} f dx = 0$	Force of order r≠0
$p_s \neq p_r$ $\omega_s = \omega_r$	$f_{tang} = 1/2\cdot B_s A_r[\sin(-(p_s-p_r)x+\varphi)+$ $+\sin((\omega_s+\omega_r)t-(p_s+p_r)x-\varphi)]$	$f_{rad} \approx ... + B_sB_r[\cos(-(p_s-p_r)x+\varphi)-$ $-\cos((\omega_s+\omega_r)t-(p_s+p_r)x-\varphi)]$
	Force wave fixed to the stator + + propagating force wave along the periphery	Force wave fixed to the stator+ + propagating force wave along the periphery
$p_s = p_r$ $\omega_s \neq \omega_r$	$f_{tang} = 1/2\cdot B_s A_r[\sin((\omega_s-\omega_r)t+\varphi)+$ $+\sin((\omega_s+\omega_r)t-2px-\varphi)]$	$f_{rad} \approx ... + B_sB_r[\cos((\omega_s-\omega_r)t+\varphi)-$ $-\cos((\omega_s+\omega_r)t-2px-\varphi)]$
	$M = -lr^2\int_0^{2\pi} f dx = -\pi r^2 l B_s A_r \sin((\omega_s-\omega_r)t+\varphi)$	
	Oscillating torques. Synchronous torques out of synchronism	Force of order r=0
$p_s = p_r$ $\omega_s = \omega_r$	$f_{tang} = 1/2\cdot B_s A_r(\sin\varphi + \sin(2\omega t - 2px - \varphi))$ $M = -\pi r^2 l B_s A_r \sin\varphi$	$f_{rad} \approx ... - B_sB_r(\cos\varphi - \cos(2\omega t - 2px - \varphi))$
	Constant force + + propagating force wave along the periphery	propagating force wave along the periphery
	Constant torque in time	Force like a fundamental wave

During the examination, we must allow positive or negative value not only for the angular velocity of the fields (the stator field relative to an assumed positive direction, the rotor field relative to the rotor, and the mechanical angular velocity of the rotor also relative to the same positive direction), but also both positive and negative values for the pole pair numbers p_s and p_r, representing the direction of rotation of the respective rotating harmonic field relative to the assumed positive direction.

The bottom rows of the table show the situation when the conditions for the formation of a constant torque in time are met. This is just the condition for the formation of synchronous parasitic torque. If this is greater than the torque of the main machine considered to be free of harmonics, the main machine "remains stuck": either at zero speed (standstill) or rotating at a certain speed (not reaching the rated speed).

The rest of the table examines the different possible variants of non-fulfillment of the conditions.

Since both radial and tangential forces will be calculated as products of trigonometric functions, the same transformations always give *two force waves* of the same magnitude. In reality, of course, their sum gives the resultant force. By studying the mathematical expression of the forces, it can be discovered that the force wave with the lower order $\underline{r}$ gives the "instantaneous mean value" of the resulting wave.

In practical calculations, only the force wave with the lower order $\underline{r}$ will be used for calculation, because it is only that being able to make the yoke (and the teeth) vibrate; the other remains out of consideration. It means that in practice we never calculate the real, resultant force, but only its instantaneous mean value (see [3] Fig. 22., [5] Fig. 137.).

We have already identified the phenomenon occurring for three possible variants. Now, the last possible variant is included also in the table in order to make the table complete.

This $\omega_s=\omega_r$, $p_s \neq p_r$ newly inserted variant is of something interesting in theory; that if the angular velocities of the stator and rotor fields are exactly the same, that always occurs during run-up or break on a certain rotor speed. Then one of the force waves does not depend on time, this means it is fixed to the stator: this is either the "envelope" curve (if p_s and p_r have the same sign) or the force with double frequency (if p_s and p_r have opposite signs), and the other one is the usual traveling / rotating wave. Richter chose just this particular case to demonstrate the creation of the one-sided magnetic pull caused by two induction waves ([5] Kapitel 4. p. 200), if $p_s \pm p_r = 1$. Otherwise, in all cases, both force waves are traveling / rotating waves. A special case is the order number r=0; then the so-called "breathing" force wave cannot be interpreted as a space wave, because it only changes in time, and is constant along the circumference. In reality, a traveling wave with twice the number of poles and approximately twice the frequency also occurs with r=0, but it remains out of consideration in practice.

Regarding Figure 3 and Table 1, note that Table 1 covers the synchronous phenomena while the processes of the harmonic asynchronous machines are governed by the usual asynchronous equivalent circuit diagram.

Before closing this chapter, however, we cannot refrain from making a critical comment regarding [8]. In the derivation, [8] did not impose any restriction on the rotor quantities ("they are of arbitrary origin"). When the conditions of equality of pole number and speed are met, according to [8] (47) p. 47, a constant torque occurs over time and it is referred as *asynchronous* torque (perhaps in reference to the title of the book). We have adopted this expression in the bottom row of our Table 1 but we have called it *synchronous* parasitic torque. The resulting torque is asynchronous torque only if the rotor quantities are *not* of arbitrary origin, but are strictly *induced by the stator* (according to the equivalent circuit diagram of the asynchronous machine). If they are of any (other) origin, the interaction between the stator and the rotor cannot be other than synchronous torque.

8. Tangential force waves

Tangential harmonic force *waves* of *low* order are generated only if $\nu_b \neq \mu_a$ ($\nu_b' \neq \mu_a'$). Since these waves cancel each other out along the circumference, i.e. they are not perceived as a torque acting on the shaft, they are generally less considered in research.

The tangential harmonic force wave occurs simultaneously with the radial one. These waves exerts bending loads on the teeth both of the stator as well as of the rotor. In a normal machine design, the magnitude of these force waves does not cause mechanical problems. However, the teeth must be checked for vibration, because the (high frequency) waves can cause the teeth to bend in resonance (even break if poorly designed). For drives with demanding noise emission requirements, this issue should also be addressed; *both* stator and rotor teeth should be checked, especially during the run-up. Their frequency is of course the same as that of radial magnetic force waves. The tangential harmonic magnetic force is considerably less than the radial one, so the examination is usually omitted. But the mass of the teeth (even with the winding) is also substantially less than the mass of the yoke, so the smaller forces can still cause the teeth to vibrate; this shows that the tangential force must be dealt with the same care.

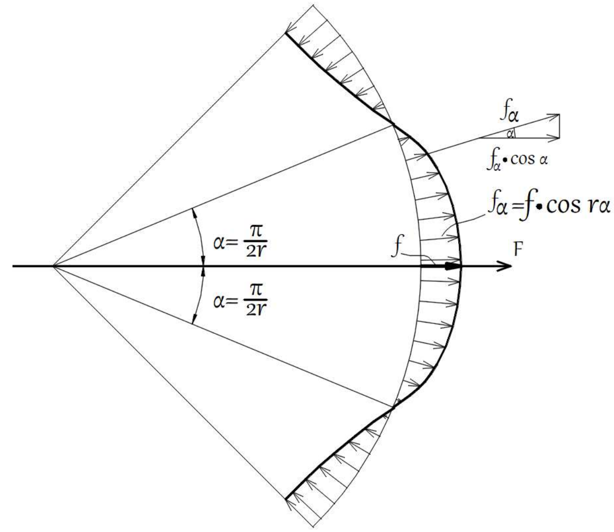

Figure 4: Radial magnetic forces (on the example of r=4)

Figure 4 shows an example of the distribution of harmonic radial forces for r=4. Here groups of sinusoidal distributed pulling and (apparently) pushing forces alternate along the periphery. It is known from physics that only tensile forces can arise in the air gap at any point of the circumference, in any operating state. The apparent contradiction has been clarified in the literature (e.g. [3]), so we will not discuss this topic now.

The tangential forces that occur simultaneously also show the same distribution (but offset in angle by π/2r), forming identical groups of forces alternately of positive and of negative direction along the periphery with respect to an assumed positive direction; the resultant torque created by them on the shaft is therefore zero.

8.1 Calculation of tangential magnetic harmonic forces

The calculation is most conveniently started from the formula for radial force [12] (26). From this, the tangential force is calculated using the well-known formula ($p\mu_a\delta/R$) regarding the ratio of tangential to radial forces [3], [7] but [8] is also suitable to apply. Let it be the maximum of a sinusoidal distributed force wave of both forces

$$f_{\tan g} = \frac{p\mu_a\delta}{R} f_{rad} = \frac{p\mu_a\delta}{R}\frac{1}{(p\tau_p l_i \delta)}\cdot(\frac{m}{2}\cdot I^2\frac{X_m}{2\pi f})\cdot(\frac{\xi_{\nu_b}\xi_{\nu_a}}{\nu_b\mu_a}\eta^2_{2\nu_a}\frac{1}{\xi_1^2}) \quad (20)$$

where p pole pairs
μ_a rotor harmonic
δ airgap
R rotor radius

After simplifications, the rotor radius still remains in the denominator. If we multiply by that in order to remove it, we get the local maximum torque

$$m_{\max} = R\cdot f_{\tan g} = \frac{1}{\tau_p l_i}\cdot(\frac{m}{2}\cdot I^2\frac{X_m}{2\pi f})\cdot(\frac{\xi_{\nu_b}\xi_{\nu_a}}{\nu_b}\eta^2_{2\nu_a}\frac{1}{\xi_1^2}) = \frac{W_{magn}}{F_{pole}}(\frac{\xi_{\nu_b}\xi_{\nu_a}}{\nu_b}\eta^2_{2\nu_a}\frac{1}{\xi_1^2}) \quad (21)$$

where F_{pole} airgap surface of one pole.

However, the formula shows a surprising fact, namely that the elementary harmonic tangential force (torque) does not depend on the rotor harmonic order μ_a that generates it (note that $\nu_b \neq \mu_a$). This is physically not possible. The explanation is that both of the cited literatures, which studied harmonic flux densities according to their subject (noise), happened to start from the harmonic induction of the *rotor* to derive the current layer of the rotor for the purpose of calculating the tangential force; hence they arrived at the well-known formula. In [5], however, the expression for the harmonic torque is derived from the harmonic induction of the *stator* for the purpose of calculating the harmonic current layer of the stator. This approach leads to the following formula

$$m_{\max} = \ldots = \frac{W_{magn}}{F_{pole}} \left(\frac{\xi_{\nu_b} \xi_{\nu_a}}{\mu_a} \eta^2_{2\nu_a} \frac{1}{\xi_1^2} \right) \tag{22}$$

which is obviously a contradiction, because $B_s{\cdot}A_r$ must be equal to $B_r{\cdot}A_s$ (Table 1).

Deriving the exact formula is laborious. The point is that the wavelengths of the stator and rotor waves are not equal. Therefore, their peaks do not coincide, but are at two different points on the circumference; the maximum of their multiplication falls between the two points. Since only the cases $r=\nu_b'\pm\nu_a'=p(\nu_b\pm\mu_a)\leq 4$ are of interest, ν_b and μ_a differ only slightly in the cases to be considered, the development of the exact formula has little practical contribution, but is of theoretical interest only. Therefore, as a compromise, the author proposes the usual geometric mean of the two harmonics. In other words, *instead of the formula described so far in the literature, the formula of the ratio is proposed to be*

$$f_{\tan g} = \frac{p\sqrt{\nu_b \mu_a}\,\delta}{R} f_{rad} \tag{23}$$

which is obviously closer to the correct theoretical value than that one used so far.

The maximum of the local harmonic magnetic torque created by the maximum tangential magnetic force wave of order <u>r</u>

$$m_{\max} = \frac{W_{magn}}{F_{pole}} \left(\frac{\xi_{\nu_b} \xi_{\nu_a}}{\sqrt{\nu_b \mu_a}} \eta^2_{2\nu_a} \frac{1}{\xi_1^2} \right) \tag{24}$$

Of course, the starting current must be substituted into (21): $I=U/X_s$, (where X_s total leakage reactance) since this is when the largest *harmonic* tangential forces are obtained. With further arrangements, we get back our original equation, [12] (13), again a cross-confirmation of our earlier formula.

Based on Fig. 4, the relative value of the torque created by a half-wave of the tangential forces of order r is proposed as

$$\frac{M_{\tan g}}{M_{breakdown}} = 2\frac{X_m}{X_s} \frac{\xi_{\nu_b} \xi_{\nu_a}}{\sqrt{\nu_b \mu_a}} \eta^2_{2\nu_a} \frac{1}{\xi_1^2} \frac{1}{2r} \frac{2}{\pi} \tag{25}$$

8.2 Consideration of magnetic conduction harmonics due to slot openings

In practice, radial magnetic force waves are usually examined at no-load and at rated load. On this slip of the motor, the magnetic conduction fluctuation harmonics, if any, play a decisive role.

However, the effect of such harmonics is different for synchronous parasitic torques, since they are naturally generated at start-up and low run-up speeds. At these times, the fundamental induction field (being in connection with said fluctuation) is about half the size of the rated one, but the currents generating the MMF harmonics are ≈5times larger, so the role of magnetic conduction fluctuation harmonics is significantly reduced and even ignored. The resulting error may be noticeable, but not so large as to affect the evaluation of the rotor slot number.

However, when considering oscillating torques (so-called torque ripple) in rated operation, which are the result of the operation of the small harmonic synchronous machines in the "out-of-synchronism" state, the magnetic conductivity fluctuation harmonics must also be taken into account (but only if the machine has an open stator slot, i.e. is designed using high-voltage technology), in the same way as for noise calculations [13] Figure 1. In this case, they are strongly dominant - also because their numerical values are not logarithmically "damped" by the dB calculation applied for noise calculations.

8.3 Pole pair numbers to be substituted in Table 1.

For all the radial force waves column and for the tangential force waves

$p_s= \nu'$, $p_r= \mu_a'$ pole pair numbers must be calculated.

For the synchronous parasitic torques

$p_s= \nu$, $p_r= \mu_a$ pole pair numbers must be calculated.

For the oscillating torques during run-up up to breakdown torque slip

$p_s= \nu$, $p_r= \mu_a$ pole pair numbers must be calculated.

For the oscillating torques around rated speed

$p_s= \nu$, $p_r= \mu_a$ pole pair numbers must be calculated but in the case of open stator slots:

the oscillating torque components that arise with the contribution of $\nu_{slot}=2mg_1q_1+1$, they must be calculated in the same way as found for radial force waves:

$p_s= \nu'$, $p_r= \mu_a'$, i.e. taking into account the amplifying effect of the magnetic conduction fluctuation wave.

Note that in the *formulae*, ν_b and μ_a must always be substituted, never ν_b' and μ_a'.

9. Conclusion

The behavior of the asynchronous motor is *determined* - for the purpose of our present investigation - by the stator and rotor *differential leakage components*.

The stator differential leakage components generate harmonic rotor currents in the rotor cage, these currents will create spatial harmonics being also differential leakage components in nature. From subject point of view, therefore, the rotor differential leakage components include all such components created by the fundamental rotor current, as well as those created by the harmonic rotor currents. At the end, each component of the stator differential leakage interacts with each component of the rotor differential leakage. This interaction will occur with any number of rotor slots. The stator and rotor components are independent of each other, and therefore the interactions are of a synchronous machine nature. The entirety of these phenomena determines the behavior of an asynchronous machine: these are the interactions that generate synchronous (harmonic) parasitic torques, radial and tangential harmonic magnetic forces.

The stator differential leakage components interact also with the harmonic rotor currents generated in the squirrel cage *by them*. Since these harmonic rotor currents depend on (are induced by) the stator differential leakage components, the phenomenon will be asynchronous in nature. These are the interactions that generate the asynchronous parasitic (harmonic) torques. The phenomenon (strongly) depends on the rotor slot number. The effect is only secondary order and cannot be measured compared to the previous one in the usual practical cases. It plays a significant role in determining the behavior of the machine, however, if $Z_2 > 1.25 \cdot Z_1$.

It is surprising, therefore, that descriptions of the generally less important small harmonic asynchronous machines appeared early on in literature, right at the beginning but descriptions of the *always* very important small harmonic synchronous machines are nowhere to be found.

10. Summary

In this paper, the completeness of MMF field harmonics in a squirrel-cage induction machine is presented, by defining a single equivalent circuit diagram containing all the harmonics and all their connections as a map, and then by defining a single table containing all the terms of all the MMF harmonics and their relations including the explanation, generation and calculation of all possible harmonic magnetic forces and torques.

In doing so, the existing definition in the literature have been verified, according to which a squirrel-cage induction machine can be modelled with a harmonic-free "main machine" and a series of "small" induction machines and "small" synchronous machines with shaft connection. The small asynchronous machines represent the asynchronous parasitic torques arising in the squirrel cage machine. The small synchronous machines represent all the magnetic forces (both radial and tangential forces incl. synchronous parasitic torques) arising in the squirrel cage machine. The small synchronous machines with the same number of poles on the stator and the rotor represent the r=0 case and the synchronous parasitic torques, and those with different numbers of poles represent the r≠0 case.

A proposal is made hereby for the general use of adjective "harmonic" for all terms regarding the phenomenon in an equivalent manner instead of "parasitic", "small" etc.

The definition of small/harmonic synchronous machines has been missing in the literature, therefore, *the definition /specification of small harmonic synchronous machines is a fundamental development* of this paper. The known but incomplete equivalent circuit diagram has thus been completed.

The completed equivalent circuit diagram also *visually* displays all harmonics and shows the complete physical interrelationship, thus greatly facilitating the understanding of the phenomena and therefore may be essential for *education*.

In practical research, each of these areas is often examined separately. However, our table organizes them in a system based on their physical relationships, and thus highlights physical relationships between the sub-areas that are often overlooked or forgotten, such as the physical relationship between oscillating torques vs. torque ripple, and between oscillating torques vs. tangential forces (occurring simultaneously with radial forces).

Based on the harmonic equivalent circuit diagram, the correctness of the formulae derived in our previous studies have been demonstrated.

Finally, the formula for tangential magnetic forces was derived. Thus, *the designer has now all the necessary relevant formulae* (synchronous and asynchronous parasitic torques, radial and tangential magnetic forces). During the derivation, it is shown that the known and long-established formula for the ratio of tangential to radial magnetic harmonic forces is not correct in principle; a proposal is given for correction.

Based on the real response of a cage rotor to stator harmonics, depending on the number of rotor slots, relatively few MMF field harmonics influence in merit the behavior of the asynchronous machine.

Since the differential leakage components are by definition spatial harmonic components of the MMF, it is proposed to designate the whole topic as the *effect of the differential leakage components* on the behavior of the squirrel cage machine.

Our system wants to serve as a guide for verifying the results of calculations using more advanced methods and for determining the direction of further research.

In this paper, together with our previous published results and formulae, we present a complete description of a new approach, which allows to compute the behavior of any asynchronous machine with squirrel cage rotors with small apparatus but still with considerable accuracy. Moreover, this *does not require any information about the machine other than the number of poles and the number of slots* (and only whether or not there is a chording on the stator), still facilitating the ultimate goal of correctly selecting the rotor slot number, on a very general way.

Acknowledgment

The author gratefully acknowledges the contributions of **István Laczkó**, Chief Engineer of the United Electric Machine Factory, EVIG, Budapest, Hungary, for his valuable comments.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1]. M.M. Liwschitz, "Field Harmonics in Induction Motors," *Transactions on Electrical Machinery*, **61(0):797–803**, 1942.

[2]. H. Möller, "Über die Drehmomente beim Anlauf von Drehstrommotoren mit Käfigankern," *Archiv für Elektrotechnik*, **18(7):400–424**, 1930. DOI: 10.1007/BF01656359.

[3]. H. Jordan, "Der geräuscharme Elektromotor," *Verlag W. Girardet*, Essen, 1950.

[4]. W. Nürnberg, "Die Asynchronmaschine," *Springer Verlag*, 1952, p. 299.

[5]. R. Richter, "Elektrische Maschinen, Vierter Band, Die Induktionsmaschine," *Springer Verlag*, 1936, 1950.

[6]. B. Heller, B. Hamata, "Harmonic Field Effects in Induction Machines," *Elsevier Scientific Publishing Company*, Amsterdam, Oxford, New York, 1977.

[7]. P.L. Tímár, "Noise and Vibration of Electrical Machines," *Elsevier*, Amsterdam, 1990.

[8]. H. Jordan, V. Klima, K.P. Kovács, "Asynchronmaschinen," *Akadémiai Kiadó* (Verlag der Ung. Akad. d. Wiss.), Budapest, Hungary, und Friedr. Vieweg und Sohn Verlagsgesellschaft GmbH, Braunschweig, 1975.

[9]. G. Kovács, "Influence of the Rotor Slot Numbers on the Parasitic Torques and the Radial Magnetic Forces of the Squirrel Cage Induction Motor; an Analytic Approach," *Proceedings of International Conference on Electrical Machines (ICEM 2022)*, Valencia, Spain, 1320–1326, 2022. DOI: 10.1109/ICEM51905.2022.9910809.

[10]. G. Kovács, "Harmonics in the Squirrel Cage Induction Motor; Analytic Calculation; Part I.: Differential Leakage, Attenuation, Asynchronous Parasitic Torques," *CES Transactions on Electrical Machines and Systems*, **7**(3)**:**320–329, September 2023. DOI: 10.30941/CESTEMS.2023.00034.

[11]. G. Kovács, "Selection of Rotor Slot Numbers in 3-phase and 5-phase Squirrel Cage Induction Motor; Analytic Calculation," *ASTES Journal*, **10**(1):60–74, 2025. DOI: 10.25046/aj100107.

[12]. G. Kovács, "Calculation of Synchronous Torques and Radial Magnetic Forces for Pole-changing Winding Using the 3//Y / 3//Y Method," *Proceedings of International Conference on Electrical Machines (ICEM 2020)*, Göteborg, Sweden, 90–96, 2020. DOI: 10.1109/ICEM49940.2020.9270765.

[13]. G. Kovács, "Harmonics in the Squirrel Cage Induction Motor; Analytic Calculation; Part II.: Synchronous Parasitic Torques, Radial Magnetic Forces," *CES Transactions on Electrical Machines and Systems*, **7**(4):404–421, December 2023. DOI: 10.30941/CESTEMS.2023.00035.

Advances in Science, Technology and Engineering Systems Journal Vol.10, No. 4, 32-40 (2025)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Integration and Innovation of a Micro-Topic-Pedagogy Teaching Model under the New Engineering Education Paradigm

Li Weiguo, Chen Yanhong*, Yang Libing, Yang Liyan

School of Land Science and Spatial Planning, Hebei GEO University, Shijiazhuang, 050031, China

A R T I C L E I N F O

Article history:
Received: 15 July, 2025
Revised: 07 August, 2025
Accepted: 09 August, 2025
Online: 25 August, 2025

Keywords:
Micro-topic teaching
Teaching model integration
Practical skills
Teaching effectiveness evaluation

A B S T R A C T

The rapid evolution of global technologies and industrial restructuring demands innovative pedagogical approaches to foster interdisciplinary engineering expertise. This research pioneers a blended instructional framework anchored in micro-topic pedagogy under the New Engineering Education (NEE) paradigm, orchestrating case studies, heuristic scaffolding, and research-driven inquiry strategies within digitally augmented learning ecosystems. A quasi-experimental study was conducted with 132 participants (53 in the experimental cohort adopting the micro-topic framework; 79 in the control group receiving conventional instruction). Performance metrics derived from assessments, perception surveys, and behavioral analytics were evaluated through boxplot visualization. Learners in the intervention cohort exhibited marked enhancements in applied skill development (mean score gain: 5.2 points, p<0.01) and active participation (85% endorsement rate). Real-time simulation tools and adaptive feedback mechanisms further amplified interdomain troubleshooting proficiency by 23%. By synergizing technology-mediated experiential learning with industry-aligned challenges, this model bridges academic rigor and professional demands, providing a scalable blueprint for engineering education innovation.

1. Introduction

The rapid advancement of global technologies and industrial restructuring has rendered traditional engineering education inadequate in addressing the demands of emerging technologies and complex engineering challenges for interdisciplinary talent [1]. In response, China's New Engineering Education (NEE) initiative has been proposed to cultivate engineers with cross-disciplinary perspectives, innovative thinking, and practical competencies [2]. This paradigm shift not only modernizes conventional pedagogical frameworks but also aligns with industrial upgrading and global competitiveness, equipping graduates with problem-solving agility in dynamic technological landscapes. Under the NEE framework, engineering education must prioritize innovation and practicality, transcending knowledge-centric approaches to emphasize hands-on skills, collaborative aptitude, and creative problem-solving. Consequently, pedagogical reforms that integrate authentic engineering contexts into curricula have become a cornerstone of higher education transformation [3,4].

Traditional engineering education, characterized by teacher-centered knowledge transmission and theoretical emphasis, fails to equip students with practical experience or interdisciplinary integration skills, hindering their adaptability to complex engineering projects [5]. Furthermore, the exponential growth of technologies — such as artificial intelligence and big data — demands rapid knowledge renewal and cross-domain collaboration, exposing the limitations of conventional pedagogies [6-8]. Universities must urgently innovate teaching models to transcend classroom boundaries and meet the NEE's broad requirements for talent development [9].

Amidst global technological acceleration and industrial transformation, engineering education is undergoing a critical shift from traditional instruction to the New Engineering Education (NEE) paradigm. Conventional models, which focus on single-discipline knowledge transfer and basic technical training, fall short in nurturing the interdisciplinary, innovative, and practical competencies required for modern engineering challenges. Recent studies across disciplines have explored blended pedagogies — such as case-based learning, inquiry-driven methods, and micro-lectures — demonstrating their efficacy in enhancing knowledge

* Corresponding Author: Chen Yanhong, School of Land Science and Spatial Planning, Hebei GEO University, China, 001528@hgu.edu.cn

https://dx.doi.org/10.25046/aj100404

retention, practical skills, and learner autonomy. For instance, a "virtual-physical integration" approach in fermentation engineering courses utilized virtual simulations and real-world case studies to deepen students applied understanding in near-authentic environments [10]. Similarly, Micro-lectures combined with case-based teaching have been shown to significantly improve self-directed learning and satisfaction in orthopedic nursing education [11]. Case-based teaching has also proven effective in medical education; for instance, case studies in geriatric Chinese medicine graduate programs not only strengthened theoretical knowledge but also significantly enhanced practical skills, innovative thinking, and self-assessment confidence [12]. In art education, inquiry-based learning fostered creativity and adaptability by guiding students to independently explore solutions to complex problems [13].

Additionally, heuristic and inquiry-based pedagogies have been shown to enhance disciplinary mastery and holistic competencies across fields. Case-based methods in ophthalmology courses have been shown to improve both student satisfaction and professional knowledge acquisition, highlighting their potential in cultivating "excellence-oriented clinicians" [14]. In materials science laboratories, inquiry-driven approaches have been revealed to significantly boost innovation capabilities, teamwork, and self-regulated learning [15]. Likewise, the integration of micro-lectures with problem-based learning (PBL) in burn surgery education has been shown to lead to measurable improvements in academic performance and learner satisfaction [16].

However, existing research predominantly focuses on optimizing isolated teaching methods within single disciplines [17-19], with limited exploration of multidimensional pedagogical integration. While methods like inquiry-based learning, case studies, and micro-lectures improve discipline-specific outcomes, they inadequately address the NEE ’ s demands for cross-disciplinary synergy, practice-oriented training, and systemic innovation. For example, Rigid skill-training frameworks in traditional models have been identified as hindering students ability to flexibly apply knowledge in complex scenarios [20].

Under this backdrop, the micro-topic pedagogy emerges as a transformative approach. Grounded in the principle of "problems as projects, solutions as research, and outcomes as achievements," this model deconstructs complex engineering challenges into manageable micro-tasks, enabling students to conduct autonomous research and solve real-world problems within constrained timelines. It synergizes three methodologies: 1) Case-based learning to contextualize theoretical knowledge in authentic scenarios; 2) Heuristic scaffolding to stimulate divergent thinking and multidimensional problem-solving; 3) Inquiry-driven exploration to foster innovation through self-directed experimentation [21].

This tripartite integration creates a flexible, practice-oriented framework that emphasizes learning-by-doing in realistic environments. By prioritizing student agency, collaborative teamwork, and critical reflection, the model bridges academic content with industrial practice. Students gain hands-on experience in independent thinking, problem-solving, and creative innovation, thereby aligning with the NEE’s demand for adaptable, interdisciplinary competencies.

2. Theoretical Foundations of the Micro-Topic-Pedagogy

2.1. Definition and Characteristics of Micro-Topic Pedagogy

The micro-topic approach refers to a small-scale (typically 4-6 weeks), short-cycle, and problem-focused instructional project. Through well-designed, real-world-related mini research topics, students conduct independent research within a limited timeframe and propose practical solutions. These micro-topics typically address real-world applications, with a moderate scope and appropriate difficulty level, aiming to develop students' ability to solve practical problems efficiently. This teaching approach provides students with opportunities for independent inquiry and hands-on practice, fully reflecting the "problems define topics, solutions drive research, and outcomes demonstrate learning" educational philosophy.

Compared to traditional research projects, the micro-topic approach is characterized by innovation and flexibility in the following aspects: 1) Short Duration, Fast-Paced Execution – Micro-topics are designed for shorter research cycles, making them adaptable to a single semester or even a few weeks. Unlike long-term projects, micro-topics align flexibly with teaching schedules and objectives. 2) Specific Problems, Clear Objectives – Micro-topics focus on well-defined real-world problems, often closely related to students' coursework, professional fields, or future career scenarios, thereby increasing student engagement and motivation. 3) Practice-Oriented with Dynamic Feedback – Students apply learned knowledge through hands-on exploration, and instructors provide timely feedback and guidance, ensuring effective learning outcomes.

The core role of the micro-topic teaching model is to enhance student engagement and foster innovative thinking. By introducing real-world problem scenarios, students not only apply theoretical knowledge in practice but also develop critical thinking and problem-solving skills. The micro-topic approach prioritizes student-centered learning, encouraging learning by doing, where students actively engage in hands-on tasks and real-world applications. This teaching method effectively reduces passive knowledge reception and stimulates active thinking, exploration, and collaboration through well-designed problem-solving activities.

Additionally, the micro-topic teaching approach offers multiple educational benefits: 1) Constructivist Learning - Rooted in constructivist theory [22,23], students develop a deeper understanding of knowledge through self-directed inquiry, improving their independent learning and innovative thinking abilities. 2) Teamwork and Collaboration – Micro-topics are typically conducted in groups, allowing students to engage in task distribution, discussion, and cooperative problem-solving, thereby enhancing their teamwork skills. 3) Career-Relevant Learning – Many micro-topics are closely linked to real-world professional contexts, helping students understand how academic knowledge translates into practical applications.

Thus, the micro-topic teaching method is not only an innovative pedagogical approach but also a key driver in fostering independent learning, teamwork, and creative thinking. Throughout this process, students deepen their practical knowledge application and continuously enhance their problem-solving competencies through hands-on experiences.

2.2. Theoretical Foundations of the Integrated Teaching Model

The uniqueness of micro-topic pedagogy lies in its capacity to flexibly integrate multiple teaching methodologies through project-driven challenges, fostering active problem-solving and innovative thinking. By synergizing case-based, heuristic, and inquiry-driven approaches, this model establishes a multidimensional blended framework that holistically enhances educational outcomes.

The micro-topic-pedagogy teaching model is distinct in its ability to flexibly integrate multiple instructional strategies, fostering students' active engagement and innovative problem-solving skills through guided project-based learning. Specifically, this model synthesizes case-based, heuristic, and inquiry-based pedagogies, forming a "multi-dimensional integration" framework that enhances overall teaching effectiveness.

1) Case-Based Learning

Case-based learning (CBL) is an instructional approach that engages students with real-world or simulated scenarios, aiming to bridge theoretical concepts with practical applications [24]. Within the micro-topic framework, CBL provides students with authentic problem contexts, reinforcing the principle that "problems define the research topics." In designing micro-topics, instructors can select domain-specific case studies and guide students through research-driven problem-solving processes. This approach enables students to apply theoretical knowledge to practical situations, fostering a deeper and more intuitive understanding of core concepts.

By engaging in case analysis and hands-on activities, students learn to extract key issues from complex scenarios and develop solutions through iterative micro-topic research. This method not only solidifies theoretical comprehension but also enhances students' ability to address real-world challenges. The integration of CBL within micro-topic teaching situates learning in a realistic context, allowing students to develop problem-solving competencies through experiential learning.

2) Heuristic Teaching

Heuristic teaching encourages critical thinking by posing thought-provoking questions, stimulating students' creativity, and fostering an exploratory mindset [25]. Within the micro-topic framework, heuristic teaching promotes the notion that "solutions drive research" by prompting students to develop and refine strategies in response to specific challenges. Through structured questioning and guided discussions, instructors help students delineate core research problems and explore potential solutions. Thought-provoking inquiries not only cultivate curiosity but also encourage students to examine issues from multiple perspectives, leading to innovative solutions. In this approach, students must not only propose countermeasures but also validate their effectiveness through research and experimentation. The iterative process of hypothesis testing and optimization integrates the principle of "solution-driven research" into micro-topic instruction.

By embedding heuristic teaching within micro-topic learning, students are encouraged to think autonomously, maximize their creative potential, and connect research processes with real-world problem-solving. This approach fosters independent inquiry and enhances students' ability to tackle complex, interdisciplinary challenges.

3) Inquiry-Based Learning

Inquiry-based learning (IBL) emphasizes student-driven exploration and experimentation to identify and solve problems. Within the micro-topic framework, IBL facilitates the design of open-ended research tasks, encouraging students to actively engage in exploration and develop innovative solutions [26]. Micro-topic instruction, grounded in IBL, immerses students in authentic research contexts where they investigate and refine their hypotheses through experimental validation. This method shifts learning from passive knowledge acquisition to active knowledge construction, where students not only absorb information but also develop insights through iterative problem-solving.

A distinctive feature of IBL is its emphasis on the learning process itself. Within micro-topic teaching, research activities, such as conducting experiments, analyzing data, and reviewing literature—are integral to the learning journey. The outcomes of these inquiries extend beyond mere problem resolution; they contribute to students' cognitive development, analytical skills, and critical thinking capabilities. By embedding IBL into micro-topic learning, students cultivate independent research competencies and develop the ability to address complex, real-world problems through systematic inquiry.

4) Multi-Dimensional Integration in the Micro-Topic Teaching Model

The core strength of the micro-topic-pedagogy teaching model lies in its ability to integrate case-based, heuristic, and inquiry-based pedagogies into a cohesive and adaptable instructional framework, Figure 1 illustrates the integration of case-based, heuristic, and inquiry-based pedagogies within the MTP framework, highlighting their roles in problem identification, solution development, and outcome validation. This multi-dimensional integration fosters an engaging, practice-oriented learning environment that enhances students innovation capacity, self-directed learning, and practical skills.

By structuring learning around specific, real-world problems, the micro-topic model enables students to engage in research-driven inquiry. Initially, problem-based cases ground their investigations in authentic contexts, followed by heuristic strategies that stimulate critical thinking and guided exploration. Finally, inquiry-based activities allow students to validate and refine their proposed solutions through systematic research. This sequential and integrative approach transforms students from passive recipients of knowledge into active explorers and problem solvers.

The micro-topic model effectively harnesses multi-dimensional integration to cultivate student agency and enthusiasm for learning. Through iterative problem-solving, students refine their approaches and develop solutions that bridge theoretical understanding with practical application. This comprehensive pedagogical strategy not only strengthens students' theoretical foundations but also enhances their cognitive flexibility and creative problem-solving abilities, preparing them to meet the

Design questions and guidance evolving demands of modern engineering education.

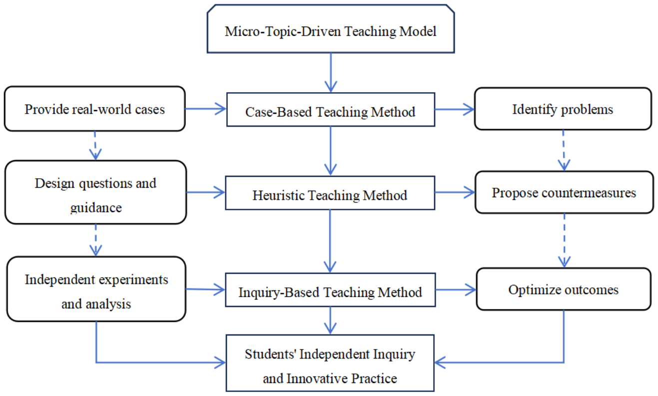

Figure 1: Framework of the Micro-Topic-Pedagogy

3. Design and Implementation of the Micro-Topic-Pedagogy Teaching Model

3.1. Principles of Micro-Topic Design and Implementation

Topic Selection Criteria: The design of micro-topics should align closely with the instructional objectives of the course, taking into account students' disciplinary backgrounds, interests, and future practical applications. Effective topic selection ensures that students not only deepen their understanding of course content but also identify real-world applications of theoretical concepts. Furthermore, selected topics should be sufficiently challenging to stimulate intellectual engagement without exceeding students' cognitive and technical capabilities. Topic selection should ensure diversity and representativeness by incorporating a range of real-world problems from different industry sectors. The design process must integrate students' academic progression, cognitive levels, and practical needs to enhance both conceptual comprehension and applied learning.

Establishing Instructional Objectives: Clearly defined and specific instructional objectives are essential to ensuring that micro-topics align seamlessly with the course curriculum. Each micro-topic should facilitate mastery of core theoretical knowledge while simultaneously fostering students' ability to solve practical problems. Thus, instructional goals should encompass not only knowledge acquisition but also the development of practical skills, teamwork, and innovative thinking. When formulating objectives, instructors should consider both the expected learning outcomes and the experiential gains students accumulate throughout the micro-topic research process.

The micro-topic implementation process follows a structured sequence:

- Preliminary Investigation: Collect relevant data to inform instructional design, ensuring alignment with both academic and industry requirements.
- Instructional Design: Based on the investigation findings, define clear learning objectives and select appropriate teaching methodologies.
- Teaching Implementation: Conduct instructional sessions while documenting classroom activities and student engagement for subsequent analysis.
- Evaluation and Optimization: Assess learning outcomes through data-driven analysis and refine teaching strategies based on student feedback.

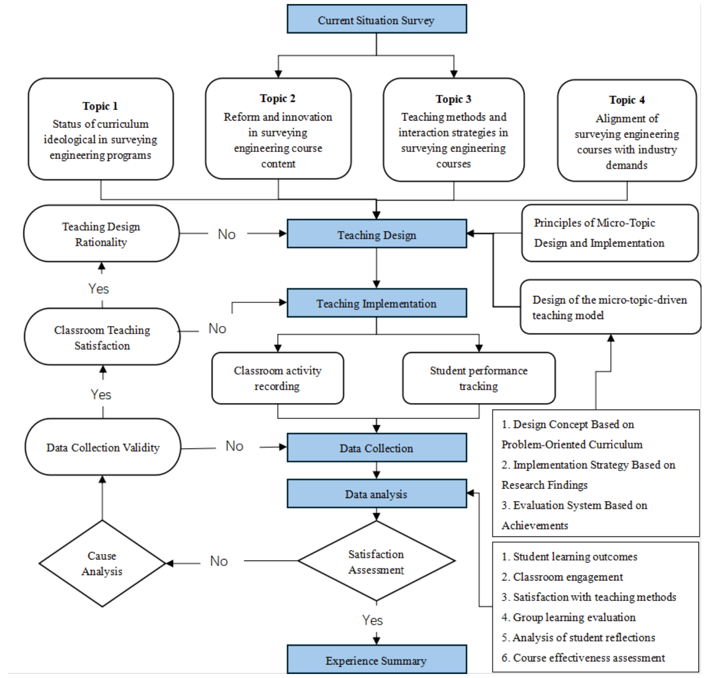

Figure 2: Design and Implementation Process of the Micro-Topic-Pedagogy Teaching Model

Taking the surveying and mapping engineering discipline as an example, the implementation process (as illustrated in Figure 2) begins with an assessment of the current state of professional courses, including ideological-political integration, curriculum development, teaching methodologies, and industry demands. This information informs the instructional design to ensure that course content remains relevant to professional industry standards.

During the teaching implementation phase, instructors systematically document student engagement and classroom interactions. Learning outcomes are then evaluated through data analysis, and iterative refinements to teaching methods are made based on feedback. This cyclical optimization process enhances teaching quality and fosters improved student learning experiences. Specifically, for surveying and mapping engineering students, this structured implementation framework provides robust support for the development of practical skills and innovative problem-solving capabilities.

3.2. Classroom Activity Design Under the Micro-Topic-Pedagogy Approach

1) Case Analysis and Discussion

Through case-based learning, students gain a deeper understanding of the background and complexity of real-world problems. In micro-topic-pedagogy classrooms, instructors should select representative real-world cases tailored to students' learning needs. These cases serve as tools to help students identify key issues, analyze root causes, and develop feasible solutions.

Case analysis not only facilitates classroom discussions but also provides students with an effective means of applying theoretical knowledge to practical scenarios.

2) Heuristic Question Design

Instructors should create thought-provoking problem scenarios that stimulate critical thinking. Under heuristic teaching principles, carefully crafted, challenging questions guide students to explore problems from multiple perspectives and investigate possible solutions. The design of heuristic questions should be closely aligned with real-world challenges, encouraging students to go beyond surface-level observations and explore underlying causes and fundamental principles. This heuristic guidance expands students' thinking and fosters an innovative mindset.

3) Inquiry-Based Task Assignments

The design of inquiry-based tasks is crucial in micro-topic teaching. Instructors should introduce tasks that are challenging, open-ended, and practice-oriented, allowing students to engage in independent exploration while solving problems. For example, in project-based learning, instructors should incorporate multiple stages of exploration characterized by variability and uncertainty. Students are encouraged to utilize experimental methods, field investigations, and data analysis to independently identify and address research problems. Inquiry-based tasks cultivate students' abilities in independent thinking and research while reinforcing their theoretical knowledge through practical applications.

3.3. Role Positioning of Teachers and Students

1) The Guiding Role of Teachers

In micro-topic-pedagogy teaching, the role of the teacher extends beyond that of a knowledge transmitter to that of a facilitator and feedback provider. Throughout the design and implementation of micro-topics, instructors are responsible for offering effective guidance in topic selection, research direction, access to learning resources, and feedback on students' findings. Teachers should encourage students to pose challenging questions, engage in reflective learning, and iteratively refine their research approaches. Additionally, timely and specific feedback is crucial in helping students refine their problem-solving strategies and enhance their learning outcomes.

2) The Active Role of Students

The micro-topic-pedagogy teaching model emphasizes student-centered learning, requiring students to actively participate in topic design, research, discussion, and presentation. In this process, students must develop both independent thinking and collaborative teamwork skills. By engaging in self-directed inquiry, they take ownership of problem-solving and progressively enhance their innovative and practical competencies. Through active participation in micro-topics, students transition from passive recipients of knowledge to proactive learners and researchers.

3) Interactive Mechanism Between Teachers and Students

Effective teacher-student interaction is fundamental to optimizing learning outcomes in micro-topic-pedagogy teaching. Teachers not only serve as knowledge facilitators but also as cognitive mentors and research advisors. By designing thought-provoking questions and structuring problem-solving tasks, instructors guide students through the entire process — from identifying research topics and formulating strategies to executing tasks and conducting reflective analysis. In this dynamic learning environment, students engage in independent inquiry and collaborative problem-solving, progressively refining their critical thinking and problem-solving abilities.

This interactive mechanism fosters a positive learning attitude and creates an iterative feedback loop that supports continuous learning and improvement. By establishing an interactive and inquiry-driven classroom environment, micro-topic teaching significantly enhances student engagement, practical application skills, and overall learning effectiveness.

4. Evaluation and Feedback on the Effectiveness of Micro-Topic Teaching

4.1. Data Analysis and Assessment Framework of Teaching Effectiveness

To objectively and systematically evaluate the effectiveness of the micro-topic-pedagogy teaching model, it is essential to establish a structured data analysis and assessment framework. First, quantitative data are collected through final exams and questionnaire surveys to measure students' overall satisfaction with the micro-topic approach, their recognition of the teaching model, and their learning experience.

After implementing the micro-topic teaching model among students in the Class of 2021, significant improvements in academic performance were observed. A quasi-experimental design was implemented with 132 students (53 in the experimental group using the micro-topic model; 79 in the control group with traditional instruction). The average score increased to 77.5, marking an approximate 5-point improvement compared to the Class of 2020 ($p<0.01$), as illustrated in Figure 3. This enhancement can be attributed to the integration of personalized learning, the combination of theoretical and practical applications, and the cultivation of critical thinking skills.

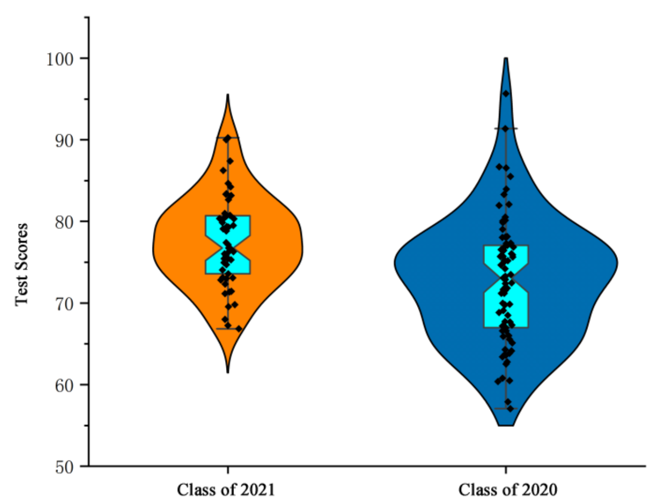

Figure 3: Comparison of Test Scores Between Class of 2021 and Class of 2020

Boxplot analysis indicates that the distribution of scores among the 2021 cohort is more concentrated, with a higher median compared to the 2020 cohort. The narrower interquartile range suggests greater consistency and stability in academic performance. Additionally, violin plot analysis reveals that the middle section of the 2021 cohort's distribution is wider, indicating a concentration of scores around the mean. In contrast, the 2020 cohort exhibits a wider distribution in the lower score range, suggesting a prevalence of lower academic performance.

These observed improvements not only reflect an overall enhancement in student performance but also demonstrate the positive impact of the micro-topic-pedagogy teaching model in improving educational quality and student learning outcomes.

4.2. Analysis of Course Learning Outcomes

A further comparison of the achievement levels of four course objectives between the 2021 and 2020 cohorts reveals that students in the 2021 cohort demonstrated overall higher attainment across all objectives, as illustrated in Figure 4. The specific findings are as follows:

The mean attainment level for each course objective in the 2021 cohort is consistently higher than that of the 2020 cohort, indicating better knowledge acquisition and skill application among students who experienced the micro-topic-pedagogy teaching model. The distribution of attainment levels in the 2021 cohort is more concentrated, reflecting improved teaching effectiveness and greater stability in students' learning outcomes.

These results suggest that the micro-topic-pedagogy teaching model significantly enhances students' achievement of course objectives, reinforcing its positive impact on engineering education quality and student learning performance.

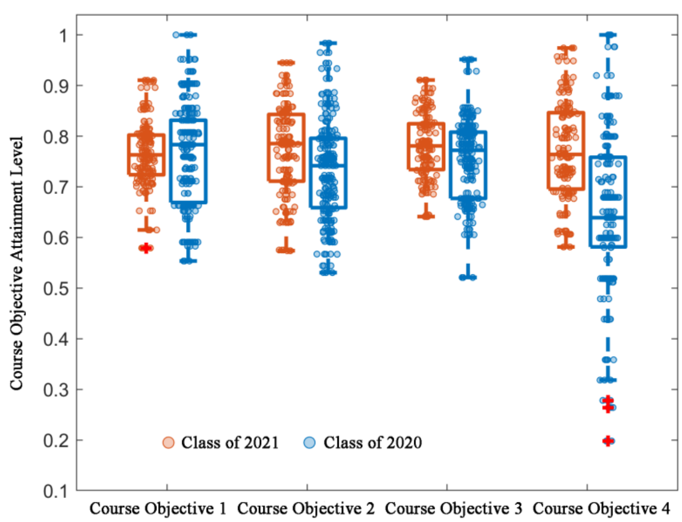

Figure 4: Comparison of Course Objective Attainment Levels

The course objectives were designed to comprehensively cover the full spectrum of learning - from theoretical understanding to practical application, and from technical proficiency to professional responsibility - aligning with the educational requirements of engineering surveying.

Course Objective 1: Students should be able to accurately articulate complex problems in engineering surveying.

Course Objective 2: Emphasizes the mastery and application of fundamental theories, requiring students to propose solutions for complex surveying challenges.

Course Objective 3: Focuses on equipping students with the ability to select appropriate technical approaches in real-world engineering scenarios.

Course Objective 4: Highlights students' ability to analyze the societal impact of engineering surveying and cultivate a sense of professional responsibility.

These well-structured course objectives provide a solid foundation for training high-quality engineering surveying professionals, ensuring that students develop both technical competencies and ethical awareness in their field.

In addition to quantitative assessments, instructors systematically recorded student's classroom performance to evaluate their engagement in team collaboration, classroom discussions, and micro-topic research activities. Student presentations, research reports, and other project-based deliverables also served as key indicators of their actual learning outcomes.

Pie chart analysis further confirmed these findings (Figure 5), showing that the attainment rates of all four course objectives exceeded 88%. This suggests that the majority of students not only successfully achieved the intended learning goals but also demonstrated high levels of enthusiasm and recognition for the classroom activities and the teaching model. The data indicate that the micro-topic-pedagogy teaching model effectively stimulates students' learning interest, while enhancing their self-directed learning abilities and problem-solving skills.

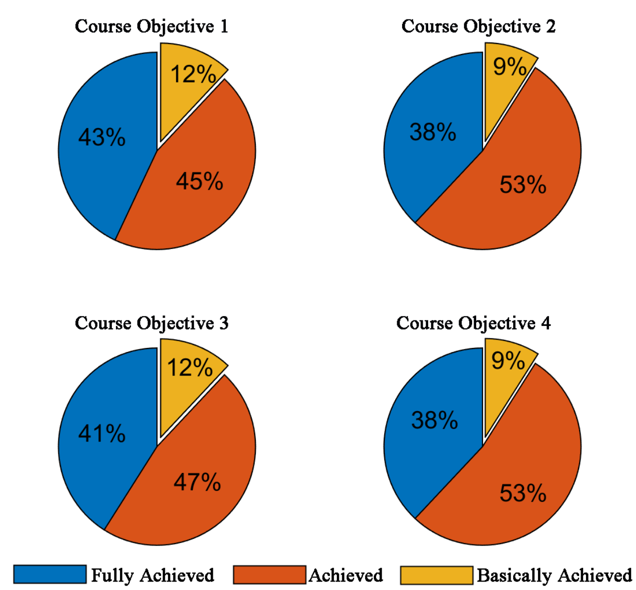

Figure 5: Achievement Levels of Course Objectives

These results highlight the strong applicability of the micro-topic teaching model in New Engineering Education (NEE) disciplines, providing a robust framework for cultivating practice-oriented and innovation-driven engineering talents.

The assessment framework is designed around the "growth as an outcome" principle, emphasizing both students' final achievements and their developmental progress throughout the research process. Beyond technical competencies, the evaluation also considers students' improvements in critical thinking, collaboration skills, and social responsibility.

By integrating boxplot and pie chart analysis, instructors can continuously refine teaching methodologies, ensuring that the micro-topic approach sustainably enhances both educational quality and student competencies. This dual-impact strategy supports the long-term advancement of teaching effectiveness and student skill development, reinforcing the micro-topic model as a valuable innovation in engineering education reform.

4.3. Student Feedback and Performance Evaluation

To comprehensively assess the impact of the micro-topic-pedagogy teaching model on students' learning outcomes and skill development, a mixed-methods evaluation framework was implemented, integrating quantitative and qualitative approaches. Quantitative analysis utilized exam scores, performance metrics, and Likert-scale surveys to measure progress in knowledge acquisition, skill application, and teamwork. Results revealed that 85% (95% CI: 80%-90%) of students rated the model as "satisfactory" or "very satisfactory," with 42% expressing strong approval (Figure 6). These findings underscore the model's efficacy in meeting diverse learning needs and enhancing educational experiences through structured, outcome-focused tasks.

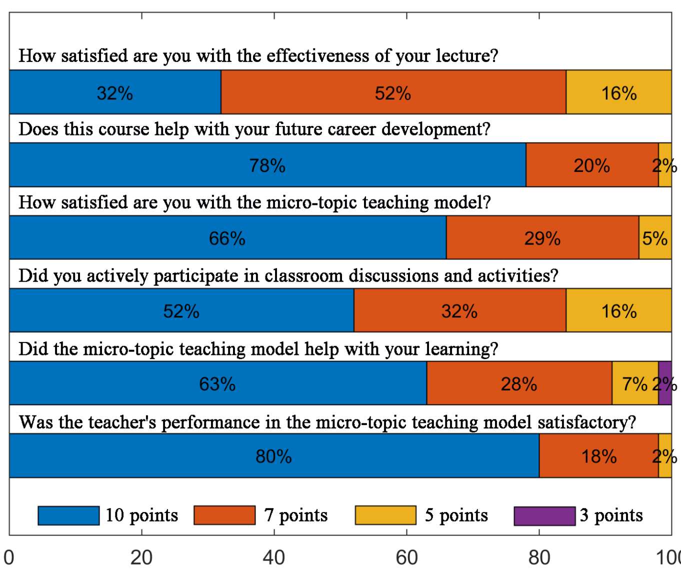

Figure 6: Student Satisfaction and Feedback on the Micro-Topic Teaching Model

Qualitative insights were derived from student reflections, classroom participation records, and team role analyses, offering nuanced perspectives on innovation capabilities, practical skills, and self-directed learning growth. This approach enabled instructors to track individual development trajectories and tailor guidance effectively. Complementing this, a dynamic assessment mechanism provided real-time monitoring of student progress through iterative feedback loops. For instance, 78% of students reported that instructor interventions significantly advanced their research, while 67% noted improved innovative thinking, highlighting the value of adaptive teaching strategies in fostering intellectual agility.

The integration of continuous feedback allowed educators to refine instructional content and pedagogical tactics iteratively. By aligning teaching adjustments with student-reported challenges and successes—such as optimizing project timelines or integrating peer-review cycles—the model evolved to better support problem-solving proficiency and critical thinking. This responsive framework not only sustained student motivation but also ensured that micro-topics remained aligned with both academic objectives and real-world relevance, ultimately cultivating adaptable, solution-oriented learners prepared for complex professional environments.

5. Advantages and Challenges of the Micro-Topic-Pedagogy Teaching Model

5.1 Advantages of the Teaching Model

One of the key advantages of the micro-topic-pedagogy teaching model is its ability to enhance student agency and engagement. By integrating real-world problems into the classroom, this model encourages students to actively participate in topic selection, research, and outcome presentation. Unlike traditional passive learning methods, the micro-topic approach transforms students into active inquirers, fostering creativity and critical thinking throughout the problem-solving process. Students are not merely receiving knowledge but rather developing a comprehensive skill set through research-based learning and real-world application. This includes teamwork, leadership, and self-directed learning abilities.

The blended pedagogical approach of case-based, heuristic, and inquiry-based learning significantly enhances teaching effectiveness, as evidenced by a 23% improvement in cross-disciplinary problem-solving competencies ($p<0.01$). As students engage in problem identification, strategy development, and result presentation, they experience an iterative learning cycle that embodies the principles of 'problems define research, solutions drive inquiry, and outcomes demonstrate learning'. The seamless integration of multiple teaching strategies not only reinforces theoretical knowledge but also strengthens students' problem-solving skills through real-world application. For example, experimental group students demonstrated statistically significant advancements in resolving interdisciplinary challenges, such as optimizing smart city infrastructure through IoT-enabled simulations. This approach effectively promotes both practical competency and innovative thinking, fostering students' comprehensive professional development.

5.2 Challenges and Recommendations for Improvement

Despite its clear benefits, the implementation of the micro-topic-pedagogy teaching model presents several challenges. First, it imposes higher instructional demands on teachers, requiring meticulous design of research topics aligned with course objectives, continuous guidance and feedback, and effective management of multiple student groups at varying research stages within limited class time. Balancing topic complexity with time constraints further complicates classroom dynamics. Second, ensuring the practical relevance and feasibility of micro-topics remains critical. Overly simplistic topics risk reducing engagement, while overly complex ones may exceed students' capabilities, leading to frustration. Striking a balance between challenge and achievability is essential, alongside aligning topics with professional aspirations and industry demands to enhance authenticity and applicability.

To address these challenges, strategic improvements are proposed. For instructional workload, group-based assignments and staged feedback mechanisms can optimize classroom management, while efficient resource allocation ensures quality guidance. For topic design, educators must calibrate difficulty levels to match student competencies and integrate real-world engineering applications to bolster relevance. For example, incorporating industry-aligned projects (e.g., sustainable infrastructure design) ensures students gain practical skills while solving authentic problems. These adjustments not only mitigate implementation hurdles but also enhance the model's educational impact.

In conclusion, the micro-topic-pedagogy teaching model significantly enhances student engagement, critical thinking, and learning outcomes. However, its effectiveness hinges on resolving challenges related to teacher workload, topic complexity, and classroom logistics. By refining topic design, optimizing resource distribution, and aligning objectives with industry needs, educators can maximize the model's potential. This approach not only supports engineering education reform but also equips students with interdisciplinary problem-solving skills essential for modern technological landscapes. Ongoing refinement through multi-stakeholder engagement will further solidify its role in fostering adaptable, innovative professionals.

6. Conclusion and Future Prospects

This study provides an in-depth exploration of the innovative micro-topic-pedagogy teaching model and its effectiveness within the New Engineering Education (NEE) paradigm. By seamlessly integrating case-based, heuristic, and inquiry-based teaching strategies, the micro-topic approach offers students a balanced platform for theoretical learning and practical application. It significantly enhances their self-directed learning abilities, creative thinking, and hands-on skills. Throughout the micro-topic learning cycle, students experience a complete inquiry and problem-solving process, from problem identification and strategy development to final presentation of results. This model not only improves academic performance but also strengthens students' competence in tackling complex engineering challenges.

The micro-topic-pedagogy teaching model presents new perspectives and practical insights for educational reform. By bridging theoretical learning with real-world problem-solving, it overcomes limitations of traditional instruction and significantly enhances students' comprehensive skills. The model proves particularly effective in developing interdisciplinary thinking, teamwork, and innovation awareness. Its flexibility and applicability make it a valuable reference model for curriculum reform across various disciplines, offering new pathways for talent development in modern higher education.

Despite its proven educational benefits, the implementation of the micro-topic approach still faces challenges. Future research and practice should further refine micro-topic design and execution. For instance, leveraging artificial intelligence (AI) and big data analytics could enable more personalized feedback and topic customization, better accommodating students' diverse learning needs. Additionally, as the demand for interdisciplinary knowledge continues to grow, integrating broader subject areas into micro-topic research will be a key direction for future exploration. Technology-driven optimization and improved resource allocation will allow the micro-topic model to adapt to increasingly complex and diverse educational environments.

Further innovation can also arise from combining the micro-topic model with other advanced teaching methodologies, such as Project-Based Learning (PBL) and the flipped classroom. The synergy of multiple pedagogical approaches would broaden learning experiences, address varied learning preferences, and further enhance instructional effectiveness. As education reform is a continuous process, the ongoing development and innovation of the micro-topic teaching model will provide new opportunities for future teaching methodologies, driving higher education toward greater innovation and practical excellence.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research was supported by the Research and Practice Project of Higher Education Teaching Reform in Hebei Province under Grant No. 2023GJJG295. The authors would like to express their sincere gratitude to the funding body for its valuable financial support, which has been instrumental in facilitating the completion of this study.

References

[1] J. Gao, G. Chen, Q. Wang, Y. Liu, "Industry-specific transformation and upgrading of surveying and mapping engineering programs under the new engineering paradigm." Bulletin of Surveying and Mapping, **2022**(5), 166–169, 2022. doi: 10.13474/j.cnki.11-2246.2022.0160

[2] M. Ye, Y. Deng, Y. Zhang, L. Zhu, "Chinese Paradigm in the Transformation of Engineering Education: Explorational Research and Theoretical Formation of 'Emerging Engineering Education'." Research on Science and Education Development, **3**(3), 18–35, 2023. doi: 10.20105/j.cnki.jstes.2023.03.003

[3] Q. Liu, "Training of innovative applied talents in materials science under the new engineering paradigm—A review of Material Technology and Equipment." Nonferrous Metals Engineering, **12**(1), 144, 2022. doi: 10.3969/j.issn.2095-1744.2022.01.021

[4] Z. Xu, Y. Li, Y. Dong, W. Zhou, "Semi-open project-based training for complex engineering problem-solving abilities." Computer Education, **2019**(2), 37–40, 2019. doi: 10.16512/j.cnki.jsjjy.2019.02.010

[5] S. Liu, "Reconstruction and innovation of talent training mode for new engineering practice teaching." Journal of Hubei Normal University (Philosophy and Social Sciences Edition), **43**(2), 113–117, 2023. doi: 10.3969/j.issn.2096-3130.2023.02.017

[6] Z. Li, J. Gao, "Exploration and practice for curriculum ideology and politics construction of characteristic specialty in the age of artificial intelligence: Taking the surveying and mapping engineering in China University of Mining and Technology as an example." Bulletin of Surveying and Mapping, **2022**(S1), 17–20, 2022. doi: 10.13474/j.cnki.11-2246.2022.0504

[7] B. Zhang, L. Xu, J. Pan, H. Li, "Development of an industry-education-integrated AI talent training system under the new engineering paradigm." Computer Education, **2023**(5), 1–6, 2023. doi: 10.16512/j.cnki.jsjjy.2023.05.034

[8] S. Wang, Y. Zhang, F. Han, L. Chen, L. "Exploration and practice of talent training for the integration of bachelor's, master's and doctoral degrees in surveying and mapping engineering major under the background of smart surveying and mapping." Bulletin of Surveying and Mapping, **2024**(8), 172–176, 2024. doi: 10.13474/j.cnki.11-2246.2024.0830

[9] O. Xie, X. Niu, Z. Cao, S. Zhu, "Exploration on the Practical and Innovative Talents Training Mode in Local Colleges under the Background of 'New Engineering'." Innovation Education Research, **8**(5), 629–634, 2020. doi: 10.12677/CES.2020.85103

[10] Zhu, C., Li, N., Fu, C., Shen, L. "Application of Case-Based Teaching Method of 'Virtual and Real Combination' in Fermentation Engineering Teaching." Journal of Dezhou University, **38**(4), 45–49, 2022. doi: 10.3969/j.issn.1004-9444.2025.02.014

[11] F. Xue, Y. Zheng, H. Lin, B. Feng, "Evaluation of micro-lecture combined with case-based teaching in orthopedic clinical nursing education." Chinese Journal of Metallurgical Industrial Medicine, **40**(5), 60–64, 2023. doi: 10.13586/j.cnki.yjyx1984.2024.05.084

[12] Z. Zhao, L. Long, Z. Liu, L. Tan, F. Jia, "Application of Case-based Teaching Method in Teaching Postgraduates of TCM Geriatrics." Chinese Medical Records, **33**(6), 15–18, 2022. doi: 10.3969/j.issn.1672-2566.2024.09.027

[13] M. K. Samira, "The Role of Arts in Education: Enhancing Creativity and Critical Thinking." Research Output Journal of Education, 2024 **3**(3):66-70. doi: 10.64252/1pc28s62

[14] J. Zhang, G. Shi, Q. Chen, J. Liu, "Application of Case-Based Teaching Method in Ophthalmology Education for Clinical Medicine under the Excellence Physician Training Program." Chinese Higher Medical Education, **39**(2), 47–50, 2023. doi: 10.3969/j.issn.1002-1701.2024.08.061

[15] T. Zhai, "Application of inquiry-based teaching in the comprehensive materials experiment course." China Educational Technology Equipment, **45**(5), 88–92, 2023. doi: 10.3969/j.issn.1671-489X.2024.15.074

[16] J. Liu, "Application of micro-class combined with problem-based learning teaching method in burn surgery teaching." Chinese Journal of Burns, **42**(4), 25–28, 2023. doi: 10.3969/j.issn.2095-0616.2022.19.026

[17] D. Opanga, N. Venuste, "The effect of the use of English as a language of instruction and inquiry-based learning on biology learning in Sub-Saharan Africa secondary schools: A systematic review of the literature." African Journal of Research in Mathematics, Science and Technology Education, **26**(3), 275–285, 2022. doi: 10.1080/18117295.2022.2141961

[18] C. Wu, X. Jiang, "Inquiry-based Teaching Method Practice in Experimental Course Teaching of Synthesis and Preparation of Nanomaterials." Journal of Hubei Polytechnic Institute, **39**(2), 54–58, 2023. doi: 10.3969/j.issn.2095-4565.2023.02.012

[19] G. Neda, R. Dragana, "Problem-based and inquiry-based learning in the teaching of nature and society." Journal of Education, Society & Multiculturalism, **3**(2), 99–116., 2022. doi: 10.2478/jesm-2022-0020

[20] M. Gube, S. P. Lajoie, "Adaptive expertise and creative thinking: A Synthetic review and implications for practice." Thinking Skills and Creativity, 35, 100630, 2020. doi: 10.1016/j.tsc.2020.100630

[21] D. Zhang, Y. Yang, S. Wang, "Application of Inquiry Teaching in the Practice of Pharmacognosy Teaching." Journal of Dali University, **4**(10), 56–59, 2019. doi: 10.3969/j.issn.2096-2266.2019.10.012

[22] J. Beattie, M. Binder, H. Beks, S. Richards, "Influence of a rural longitudinal integrated clerkship on medical graduates' geographic and specialty decisions: A constructivist grounded theory study." BMC Medical Education, **24**(1), 795, 2024. doi: 10.1186/s12909-024-05793-5

[23] J. Jin, X. Duan, X. Li, W. Zhang, "Research on instructional design in the era of intelligent education from a new constructivist perspective." Journal of Beijing Electronic Science and Technology Institute, **32**(2), 116–127, 2024. doi: 10.3969/j.issn.1672-464X.2024.02.012

[24] L. Ying, "Application of case-based teaching in college students' mental health education course reform." China Educational Technology Equipment, **2022**(9), 114–116, 2022. doi: 10.3969/j.issn.1671-489X.2022.09.114

[25] C. Li, Z. Wang, L. Sun, "Practical exploration of heuristic teaching in petroleum engineering education." Education Modernization, **5**(34), 48–49, 2018. doi: 10.16541/j.cnki.2095-8420.2018.34.023

[26] R. Sam "Systematic review of inquiry-based learning: assessing impact and best practices in education." F1000Research 2024, 13:1045, 2024. doi: 10.12688/f1000research.155367.1