APPLICATION OF MACHINE LEARNING TO FORM A KNOWLEDGE BASE OF INTELLIGENT INFORMATION SYSTEMS

Abdullaeva Ozoda Safibullaevna
Namangan Institute of Engineering-Construction Doctor of Pedagogical Science (DSc), Professor
E-mail: ozoda.safibullayevna121620@gmail.com
Abdullajonova Nurzoda Nurmakhamadovna
Doctoral student at Namangan Institute of Engineering-Construction,
E-mail: nurzoda.nur27@gmail.com

https://doi.org/10.5281/zenodo.14789514

Abstract: In recent years, the integration of artificial intelligence (AI) and intelligent information systems (IIS) into various sectors has been accelerating, particularly in education. These advancements aim to improve the quality of education through automation, intelligent analysis, and dynamic decision-making. One key area where intelligent information systems have shown significant potential is in psychological diagnostics for students. These systems can play a vital role in helping assess students' abilities, provide personalized learning paths, and assist in career guidance, fostering better educational outcomes.

This article outlines the stages of developing an Intelligent Information System for Psychological Diagnostics, which aims to assess students' psychological state, cognitive abilities, and personal interests. This system will also provide the necessary data to guide students in making informed academic and career decisions. Below, we explore the stages of development and functionality of this system, focusing on the benefits it can bring to both educators and students.

Keywords: information system, intellectual system, intellectual knowledge base, database, intellectual test system

ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ФОРМИРОВАНИЯ БАЗЫ ЗНАНИЙ ИНТЕЛЛЕКТУАЛЬНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

Аннотация: В последние годы интеграция искусственного интеллекта (ИИ) и интеллектуальных информационных систем (ИИС) в различные секторы ускоряется, особенно в образовании. Эти достижения направлены на повышение качества образования за счет автоматизации, интеллектуального анализа и динамического принятия решений. Одной из ключевых областей, в которой интеллектуальные информационные системы продемонстрировали значительный потенциал, является психологическая диагностика для студентов. Эти системы могут играть важную роль в оценке способностей студентов, предоставлении персонализированных путей обучения и содействии в профориентации, способствуя лучшим образовательным результатам.

В этой статье описываются этапы разработки интеллектуальной информационной системы для психологической диагностики, которая направлена на оценку психологического состояния студентов, когнитивных способностей и личных интересов. Эта система также предоставит необходимые данные, чтобы направлять студентов в принятии обоснованных академических и карьерных решений. Ниже мы рассмотрим этапы разработки и функциональность этой системы, сосредоточившись на преимуществах, которые она может принести как преподавателям, так и студентам.

Ключевые слова: информационная система, интеллектуальная система, интеллектуальная база знаний, база данных, интеллектуальная тестовая система

INTRODUCTION

Machine learning (ML) plays a crucial role in forming and evolving the knowledge base of intelligent information systems (IIS). The use of machine learning for this purpose is highly relevant in today's data-driven world, where intelligent systems must continuously learn and adapt to changing environments and user needs. The following points highlight the key reasons why machine learning is indispensable in the formation of a knowledge base for IIS.

Machine learning algorithms are particularly effective in automatically extracting patterns, relationships, and insights from large volumes of raw, unstructured, and structured data. These patterns and relationships can then be used to form the foundational knowledge of an intelligent system. Traditional rule-based systems often rely on manually coded knowledge, which is time-consuming and limited in scope. In contrast, ML techniques, such as supervised learning, unsupervised learning, and reinforcement learning, allow intelligent systems to learn directly from data.

Machine learning algorithms enable intelligent systems to adapt over time. As new data becomes available, the system can update its knowledge base without human intervention. This ability to learn from new data allows the system to refine its decision-making processes and improve its performance, making it more dynamic and accurate in handling real-world scenarios.

The relevance of using machine learning to form a knowledge base for intelligent information systems is undeniable. Machine learning enhances the adaptability, scalability, and predictive power of these systems while enabling them to process vast amounts of data autonomously. By leveraging ML, intelligent systems can continuously update and refine their knowledge base, making them more capable, efficient, and reliable in solving complex, real-world problems. As data volumes and complexity continue to grow, machine learning will remain a fundamental technology for shaping the knowledge bases of the next generation of intelligent systems.

METHODOLOGY

Methodology for Applying Machine Learning to Build a Knowledge Base for Intelligent Information Systems

Building a knowledge base (KB) for intelligent information systems (IIS) using machine learning (ML) involves a series of structured steps designed to extract valuable knowledge from data and use it for decision making, reasoning, and problem solving. Machine learning techniques play a central role in automating the extraction and structuring of knowledge, allowing intelligent systems to adapt and improve over time.

Here is a detailed methodology for applying machine learning to build a knowledge base for an intelligent information system: creation of an intelligent information system called "Psychological Diagnostics System", the purpose of which is to determine the abilities of students, organize effective training in a certain direction, through an intelligent system, the basis of which is knowledge bases, databases. The designed information system will ensure effective work with students and their parents, intelligent data analysis, quality control [1-3].

Here's a detailed methodology for applying machine learning to create a knowledge base for an intelligent information system:

1. Data Collection and Preprocessing

The first step in applying ML to create a knowledge base is to collect relevant data from various sources. This data may include structured data (e.g., databases, spreadsheets) and unstructured data (e.g., text documents, images, sensor data).

Data preprocessing is essential to clean and prepare the data for analysis. The raw data may contain noise, missing values, duplicates, or inconsistencies that could affect the performance of ML algorithms.

2. Feature Selection and Extraction

Machine learning models rely on input features that describe the data. Feature selection is the process of identifying and selecting the most relevant features, which improves model efficiency and helps in reducing overfitting.

Filter Methods: Statistical tests (e.g., Chi-square test, ANOVA) to identify important features.

Wrapper Methods: Recursive Feature Elimination (RFE) to iteratively select the most significant features.

Embedded Methods: Lasso regression and decision trees that automatically select features during training.

3. Machine Learning Model Selection

In supervised learning, the system learns from labeled data, where the outcome (target) is known. The goal is to find patterns that map the input features to the target outcomes.

Classification Models: Decision trees, support vector machines (SVM), k-nearest neighbors (KNN), logistic regression.

Regression Models: Linear regression, random forests, gradient boosting.

These models are suitable for creating knowledge related to tasks such as classification (e.g., spam detection, sentiment analysis) or regression (e.g., predicting numerical values, time series forecasting).

Unsupervised learning is used when the data is unlabeled, and the goal is to uncover hidden structures or patterns.

Clustering: K-means, DBSCAN, hierarchical clustering.

Dimensionality Reduction: Principal Component Analysis (PCA), t-SNE.

Unsupervised learning helps in forming knowledge representations such as groupings (clusters) or latent variables that can be used in knowledge extraction and discovery tasks.

Reinforcement learning (RL) allows systems to learn by interacting with their environment and receiving feedback. This is useful for creating adaptive systems that improve decision-making over time based on reward signals.

Algorithms: Q-learning, Deep Q-Networks (DQN), Policy Gradient Methods.

RL is ideal for applications where the knowledge base evolves through ongoing interactions (e.g., robotics, autonomous vehicles).

Deep learning models, such as neural networks, can automatically learn high-level features and are particularly useful when handling complex, unstructured data.

Convolutional Neural Networks (CNNs): For image recognition and feature extraction.

Recurrent Neural Networks (RNNs): For sequential data (e.g., time-series, text).

Deep learning allows IIS to automatically learn and generate complex knowledge representations, making it ideal for tasks like natural language processing (NLP) and image analysis.

4. Model Training and Evaluation

Once a suitable machine learning algorithm has been selected, the model is trained on the prepared data. The training process involves optimizing model parameters to minimize the error between predicted and actual outputs.

Loss Function: Measures the difference between predicted and actual outputs.

Optimization: Gradient descent and its variants (e.g., Adam, SGD) to update the model parameters.

5. Knowledge Representation and Structuring

Once the model is trained, the learned knowledge must be represented in a form that the system can use effectively for decision-making and reasoning.

For many tasks, decision trees can be used to represent knowledge explicitly, where each node corresponds to a decision or a rule based on feature values.

In cases where uncertainty exists, probabilistic models (e.g., Bayesian networks) can represent knowledge in a way that accounts for uncertainty and probabilistic reasoning.

For representing structured and interrelated knowledge, ontologies can be created to describe concepts, entities, and their relationships. Semantic networks provide a way to model relationships among knowledge elements (e.g., "A cat is an animal").

6. Knowledge Base Update and Maintenance

An intelligent information system requires its knowledge base to evolve over time to maintain relevance and accuracy. This involves continuous learning and adaptation.

Machine learning models can be updated incrementally as new data arrives, ensuring that the knowledge base stays up-to-date.

Periodic retraining with fresh data ensures that the system incorporates new patterns and insights into its knowledge base.

It’s essential to monitor for errors and inconsistencies in the knowledge base and take corrective actions through error detection methods and model re-evaluation.

The methodology for applying machine learning to create a knowledge base for intelligent information systems involves data collection, preprocessing, model selection, training, evaluation, and continuous updates. By leveraging ML techniques, intelligent systems can autonomously extract valuable knowledge from data, structure it effectively, and apply it to solve complex problems. This dynamic and adaptive approach ensures that the knowledge base evolves and remains relevant in real-time, making it an essential component of modern intelligent systems [4,5].

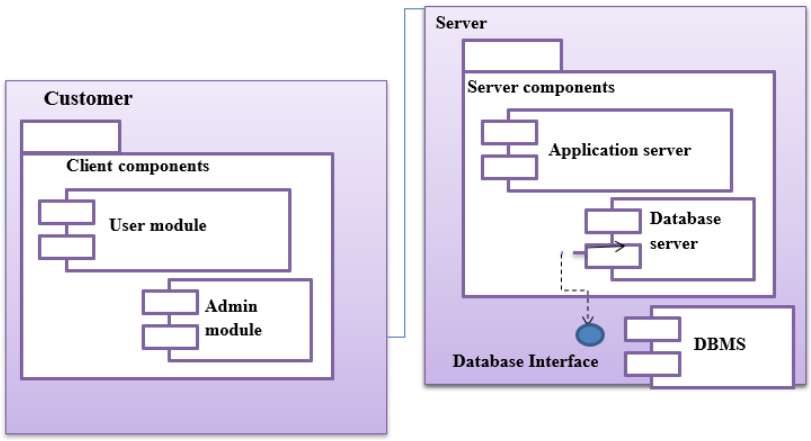

Fig.1. UML diagram for system architecture

REALIZATION OF THE CONCEPT

The following algorithms were developed to create the intellectual knowledge base of the psychological diagnosis intellectual test system:

Algorithm 1: "Scheme of User Interaction with the Intelligent System".

We divided the users of our intellectual systems into the following groups:

– admin is the technical manager of the software tool, debugger and the person fully responsible for the system operation.

– a student is a person who registers with the software, uses the information in it and creates his own database of results.

– a parent is a person who monitors the result window and can communicate with the teacher through the system if necessary.

– a teacher is a person who monitors the results window and can communicate with parents through the system if necessary (Figure 2).

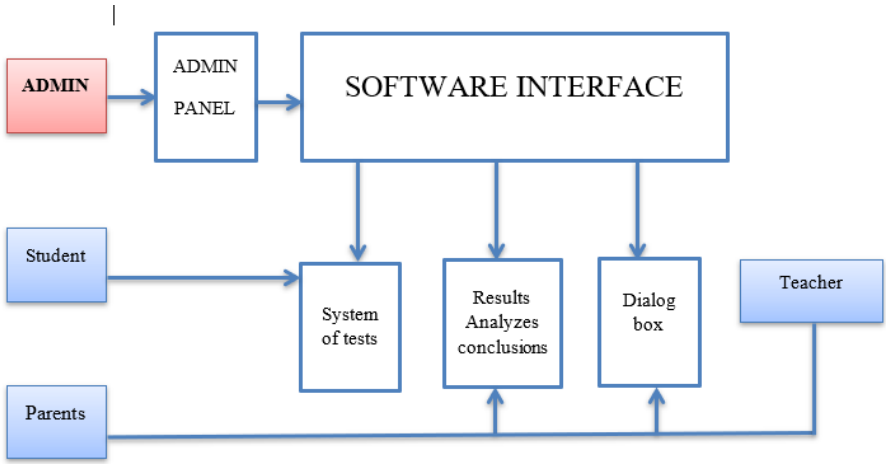

Fig.2. General scheme of interaction components

Algorithm 2: "Scheme of general functions of the knowledge base".

"Working with Knowledge Base" generally includes the following functions: "Create Knowledge Base", "Add Knowledge", "Modify Knowledge", "Remove Knowledge", "Work with Knowledge Base", " Maintaining the knowledge base. These functions are connected as follows.

Algorithm 3: "Database Schema prepared in SQLite3".

In SQLite MBBT, student (student), Category (subjects), Sub_category (science stages), Result (results), PSTSubcategory (psychological tests), PSTResult (Psychological test results), PSTQuestion (Psychological test questions), PSTAnswer (Psychological test answers), tables named Question (questions), Answer (answers) and their corresponding fields are created and connected as shown in the figure (Figure 3).

The developed software tool includes working with a database, so existing database management systems are reviewed. The most popular of them: MySQL, Oracle DB, PostgreSQL, MongoDB, SQLite Microsoft SQL server, IBM DB2 [6-9].

When selecting a database management system (DBMS) for the intellectual information system, the following factors should be considered:

1. Data requirements: The nature and volume of data that the system will handle will influence the choice of DBMS. For example, if the system deals with large amounts of unstructured data, a NoSQL database like MongoDB may be a better choice than a traditional relational database like Oracle.

2. Performance: The speed and efficiency of the DBMS will affect the overall performance of the system. Factors such as indexing, caching, and query optimization should be considered.

3. Scalability: The ability of the DBMS to handle increasing amounts of data and users over time is an important consideration.

4. Security: The security features of the DBMS, such as encryption, authentication, and access control, should be evaluated to ensure that they meet the requirements of the system.

5. Cost: The cost of licensing, maintenance, and support for the DBMS should be considered.

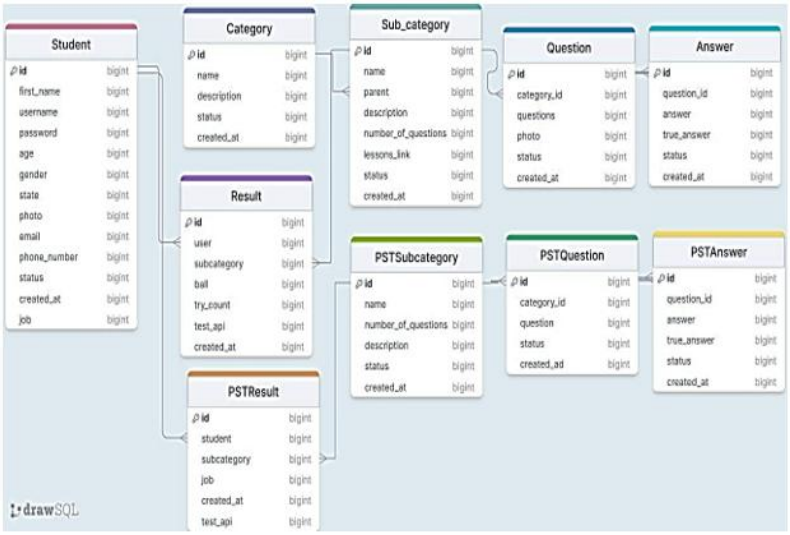

Fig.3. The scheme of connecting data with each other in the database management system (SQLite3)

Based on these factors, some of the popular DBMS options for the intellectual information system are:

1. MySQL: A popular open-source relational database management system that is known for its speed and scalability [10].

2. PostgreSQL: Another open-source relational database management system that is known for its robustness and support for advanced features like JSON data types and full-text search [11].

3. MongoDB: A NoSQL document-oriented database that is designed for handling unstructured data and offers high scalability and performance.

4. Microsoft SQL Server: A popular commercial relational database management system that offers advanced features like business intelligence and data warehousing.

5. Oracle DB: A commercial relational database management system that is known for its scalability and security features [12].

Ultimately, the choice of DBMS will depend on the specific needs and requirements of the intellectual information system.

For creating intelligent systems, we present the advantages of the following software: Python, Django, Peewee, Java Script [13,14].

DISCUSSION OF RESULTS

In this chapter, the process of designing a software tool of the issue selected for consideration of "Methods of designing and creating an intellectual information knowledge base" has been covered. In this, the construction of the model, the structure of the algorithms were mentioned, and the software tools necessary for creating a program based on the constructed algorithms were studied. After analyzing the necessary software tools for the implementation of the project, the most suitable tools were selected. The capabilities of each of the selected software tools have been explored.

The software tool is in the form of a platform, Python and Javascript programming language and Django Rest Framework, React JS libraries were used for the BackEnd part. The purpose of using this is that the API, that is, the data sent for the Frontend part of the program with the help of these libraries, will be very fast and safe, in which the development of the program is convenient compared to other programming languages and their Frameworks, which provides time saving, modern, convenient and does not require excessive time. SQLite was used for the database.

REFERENCES

1. Abdullayeva O. S. Scientific bulletin informatics 2016, issue 3 udk: 004.94: 622.794 On increasing the effectiveness of using game technologies in teaching information technologies //Scientific bulletin. - S. 67.
2. Abdullayeva O. S., Ismanova M. D. IMPROVING THE EFFICIENCY OF THE PROCESS OF PREPARATION FOR TEACHING ACTIVITIES OF THE UNIVERSITY STUDENTS //Теория и практика современной науки. – 2016. – №. 5. – С. 3-5.
3. Abdullaeva O. S. The course of study-Pedagogical Competence of the Engineer-Teacher‖ //Monograph. Publisher: LAP Lambert Akademik publishing & Co. KG, Saarbrucken, Germany. – 2018. – С. 189.
4. Abdullaeva O. S. Pedagogical competence of a professional teacher (on the example of 5330200-Informatics and information technology) //Study guide. Publisher: LAP Lambert Academic publishing & Co. KG, Saarbrucken, Germany. – 2018. – С. 105.
5. Abdullaeva O. S. Technology enhance the educational process of professional colleges and academic lyceums (for example, studying the course of informatics and information technology) //Monograph. Publisher: Navruz.-Tashkent. – 2016. – С. 189.
6. Abdullaeva O. S. Technologies of increasing the efficiency of the process of preparing for pedagogical activity of students in the direction of vocational education //Diss… PhD in pedagogical sciences: 13.00. 01.-Namangan. – 2018. – С. 156.
7. Abdullayeva O., Beknazarova S. MEDIA CULTURE OF MODERN YOUTH IN UZBEKISTAN // The current state of media education in Russia in the context of global trends: materials of the II international scientific conference. Taganrog, October 15, 2020 – DirectMedia LLC, 2020. – P. 22
8. Beknazarova S., Abdullayeva O., Abdullayev Z. METHODS OF PROCESSING INFORMATION RESOURCES IN MULTIMEDIA SYSTEMS //Конференции. – 2021.
9. Abdullaeva O.S., Beknazarova S.S. Technologii povyshenia effektivnosti vneklassnogo uchebnogo protsessa //Issledovatel, SShA. – 2016. – no. 8. - S. 30.
10. Safibullaevna, A.An integrated approach to the effectiveness of training masters of technical higher educational institutions//International Journal of Mechanical Engineeringthis link is disabled, 2022, 7(1), стр. 1170–1173
11. Safibullaevna, A.O., Inomjonovich, I.A., Murodullayevich, M.N., ...Avazovna, S.N., Muxammatrizoyevich, N.A.ORGANIZATIONS AND CONTROL OF SELF-EDUCATIONAL EDUCATION MAGISTRATORS TECHNICAL EDUCATIONAL EDUCATION//International Journal of Mechanical Engineeringthis link is disabled, 2022, 7(1), стр. 1174–1177
12. Safibullaevna, A.O., Karamtdin Kizi, J.M., Engalichev, M.I., Safibullaevna, B.S.Online-learning organization methodology as component of it technologies at students of technical universities//2020 International Conference on Information Science and Communications Technologies, ICISCT 2020, 2020, 9351487

13. O.S.Abdullayeva, A.Isomiddinov. Diagnostics and identification of conditions of effectiveness of the process of preparation for pedagogical activity of students on the direction of "professional education"// International journal of mechanical and production engineering research and development (ijmperd), issn (online): 2249-8001; issn (print): 2249-6890; impact factor(jcc) (2019): 8.8746.
14. Abdurasulovna, Khusayinova Guzal. "OPPORTUNITIES FOR THE FORMATION OF STUDENTS'CREATIVE THINKING IN TECHNOLOGY CLASSES." Galaxy International Interdisciplinary Research Journal 10.5 (2022): 187-191.
15. M Madaliev, M Usmonov, M Ismoilov, I Bilolov, S Israilov, N Abdullajonova, Khurshida Rajabova. Analysis of jet currents using the SST turbulence model. E3S Web of Conferences 538, 01019
16. X Qodirov, X Rajabova, N Abdullajonova, Z Otaxonova, I Aliev. On analytical study of heat transfer phenomenon in special-shape soldering iron. E3S Web of Conferences 508, 05006
17. Abdullajonova Nurzoda Nurmaxamadovna, Development of Critical Thinking in Media Education, Excellencia: International Multi-disciplinary Journal of Education (2994-9521) 2024, 561-563
18. A.N Nurmaxamadovna, A.O Safibullayevna. Ta’lim jarayonini rivojlantirishda qo‘llaniluvchi media ta’limining asosiy modellari - Ustozlar uchun, 2024