MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE REPUBLIC OF KAZAKHSTAN

СӘТБАЕВ УНИВЕРСИТЕТІ | SATBAYEV UNIVERSITY

GRADUATE PROFILE (DOCTOR'S DEGREE)

**Of Educational Program
8D06102 – Machine Learning & Data Science
(M094 «Information Technologies»)**

Almaty, 2024

CONTENTS

INTRODUCTION

The educational program «8D06102 Machine Learning & Data Science» has been developed within the framework of the 8D06 «Information and Communication Technologies» direction and aims to prepare highly qualified specialists capable of solving complex tasks in the fields of machine learning and data science. In the context of rapid technological advancements and the growing volume of information, graduates of this program will be able to apply their knowledge to implement innovative solutions across various sectors, including business, healthcare, finance, and others.

The program includes an in-depth study of disciplines such as «Machine Learning & Deep Learning», «Big Data Storage Systems and Computations» and «High Load Distributed Computing», as well as courses on research methods and academic writing. These disciplines provide a comprehensive approach to education, combining theoretical knowledge with practical skills for effective handling of big data, developing intelligent systems, and integrating automated analysis and data-driven decision-making processes.

A key aspect of the program is preparing graduates with both technical and analytical competencies, including critical thinking and the ability to conduct scientific research, which are crucial in a rapidly evolving technological landscape. Special attention is given to courses like «Doctoral Research Work» and «Research Practice», which help develop independent research skills and prepare students for defending their doctoral dissertation.

The learning process is organized based on a credit-based system, allowing for flexible formation of each doctoral candidate's educational trajectory. The program includes internships, participation in scientific projects, dissertation defense, and culminates in a final assessment that comprehensively evaluates the professional readiness of graduates.

Thus, the «8D06102 – Machine Learning & Data Science» program is aimed at developing competitive specialists who can design and implement advanced solutions for data analysis and machine learning, effectively integrating into the modern IT sector and scientific research.

1 Goals and Objectives of the Educational Program «8D06102 – Machine Learning & Data Science»

Objective:

The educational program aims to train scholars capable of independently conducting scientific research, developing complex software solutions, working effectively in teams, and confidently navigating modern aspects of artificial intelligence and data science. The program focuses on preparing highly qualified specialists equipped with competencies for practice-oriented work in the fields of data analysis, machine learning, and artificial intelligence, as well as for conducting original scientific research and implementing innovative solutions across various industries.

Tasks:

- Provide doctoral students with practical skills and knowledge necessary for working in the fields of data analysis, machine learning, and artificial intelligence.

- Develop the ability to apply theoretical knowledge to solve real-world problems.

- Create conditions for conducting original scientific research in the field of machine learning and data science.

- Support the publication of research findings in international and domestic peer-reviewed journals.

- Train doctoral students in research methods and scientific analysis.

- Develop skills for designing and implementing effective machine learning algorithms to solve practical problems.

- Teach doctoral students how to create and optimize algorithms for various applications and industries.

- Prepare doctoral students to design and implement complex software systems for big data analysis.

- Teach the use of modern tools and technologies utilized in software development.

- Foster the ability to work effectively in interdisciplinary teams and collaborate with other specialists.

- Train doctoral students in communication and presentation skills to effectively convey the results of their work.

- Promote lifelong learning and professional development among doctoral students.

- Ensure an understanding of the social, economic, and environmental aspects of working with data.

These tasks aim to shape graduates ready to meet the challenges and demands of the modern industry in machine learning and data science.

2 List of Qualifications and Positions

Graduates of the specialty 8D06102 – «Machine Learning & Data Science» are awarded the degree of Doctor of Philosophy (PhD) in Technical Sciences upon successful defense of their doctoral dissertation. The educational program is developed in accordance with the State Compulsory Standards of Higher and Postgraduate Education, approved by the Order of the Minister of Science and Higher Education of the Republic of Kazakhstan dated July 20, 2022, No. 2 (registered in the State Register of Regulatory Legal Acts under No. 28916). It aims to achieve the learning outcomes that define the content of curricula, including academic and individual study plans for students, as well as syllabi for courses.

Graduates of the specialty 8D06102 – «Machine Learning & Data Science» can hold the following positions:
- Lead Machine Learning Engineer;
- Senior Data Analyst;
- Team Leader for Artificial Intelligence Algorithm and Model Development;
- Big Data Architect;
- Lead Cloud Computing Engineer;
- IT and Data Science Project Manager;
- Researcher in Machine Learning and Data Analysis;
- Professor or Associate Professor in higher education institutions.

The professional sphere of graduates encompasses all sectors of the economy where methods and technologies of machine learning and data analysis are actively applied, including healthcare, finance, retail, information technology, and scientific research.

The professional objects of a Doctor of Philosophy (PhD) in this field include:
- Innovative machine learning models and data analysis algorithms designed for complex applied and scientific tasks;
- Systems for big data processing and storage, including distributed and cloud platforms;
- Specialized platforms and technologies for high-load computing and their integration with artificial intelligence systems;
- Tools and methodologies for advanced analytics and data interpretation in various interdisciplinary fields.

3 Descriptors

The requirements for the level of preparation of doctoral students under the 8D06102 – «Machine Learning & Data Science» program are based on the Dublin Descriptors for the third level of higher education (doctoral studies). They outline the necessary competencies expressed in learning outcomes. The descriptors describe the key abilities of graduates as follows:

- Possess deep and specialized knowledge in machine learning and data science, including the latest theories, methodologies, and technologies. Doctoral students should be able to utilize this knowledge to conduct original research, create innovative solutions, and develop software systems capable of addressing contemporary challenges by integrating the latest advancements in science and technology.

- Perform comprehensive analysis and interpretation of large datasets using modern methods in machine learning, statistics, and artificial intelligence. Doctoral students must have the skills to adapt algorithms and tools to diverse tasks and requirements, ensuring high accuracy, reliability, and scalability of results.

- Design and refine complex models and machine learning algorithms to address research and practical challenges. Graduates should be capable of developing new methodologies, algorithms, and tools while optimizing existing approaches to enhance data processing performance and efficiency.

- Critically evaluate and interpret research and development results, considering their scientific, social, and ethical implications. This includes understanding the environmental and societal impacts of research, adhering to standards of responsible research, and recognizing ethical issues in the creation and application of technologies.

- Demonstrate skills in knowledge dissemination and presenting complex ideas in an accessible manner to both scientific communities and broader audiences. Graduates should effectively present their research findings, publish scientific papers in peer-reviewed journals, and speak at conferences, sharing their discoveries and insights.

- Exhibit a high degree of independence in research activities, showing readiness for continuous self-learning and adaptation to rapidly changing technological demands. Graduates should demonstrate a commitment to expanding their knowledge and skills in data science and machine learning by leveraging innovative approaches and resources for ongoing professional development.

4 Competencies upon completion of training

4.1 Requirements for the Key Competencies of Graduates of Scientific and Pedagogical Doctoral Studies

A graduate of the scientific and pedagogical doctoral program in 8D06102 – «Machine Learning & Data Science» should possess a comprehensive set of knowledge, skills, and abilities that enable not only the application of advanced technologies but also contributions to their development and the effective transfer of knowledge to the next generation of specialists. The competencies formed during the program encompass the following key areas:

1) Knowledge of:

- The role of science, education, and advanced technologies in machine learning and data science in the development of modern society, and their significance in solving global issues, including sustainable development, healthcare, economics, and ecology.

- Modern approaches to the development and implementation of machine learning algorithms needed for processing and analyzing large data sets, addressing challenges in areas such as predictive analytics, natural language processing, and risk management.

- Current philosophical and methodological aspects of research in artificial intelligence and data science, including the importance of an interdisciplinary approach to gain a deeper understanding and solve complex modern problems.

- The modern challenges and new opportunities associated with the use of machine learning and data analysis in various sectors, such as finance, healthcare, energy conservation, and social sciences.

2) Knowledge of:

- Key methods and principles of organizing scientific activities, as well as essential approaches for conducting research in the areas of big data, distributed computing, and intelligent data analysis.

- The methodology of scientific cognition and the structure of scientific research, including aspects of sustainable development and innovative technologies in machine learning.

- Psychological aspects of cognitive activity necessary for enhancing the effectiveness of teaching and creating a motivating learning process.

- Modern algorithms and tools for processing large data sets, text analysis, predictive analytics, and behavioral modeling that ensure high accuracy and efficiency in analytical processes.

3) Skills to:

- Apply acquired knowledge to independently conduct scientific research and implement innovative solutions in machine learning and big data analysis.

- Critically assess and apply existing concepts, methods, and algorithms, such as deep learning, clustering, and predictive analytics, adapting them to research tasks.

- Integrate knowledge from different fields to solve complex scientific and practical problems in areas such as medical diagnostics, financial analysis, and sustainable development.

- Use interactive methods and innovative approaches, including project-based learning, data analysis from real-life examples, and critical case studies, in educational practice.

- Develop and improve machine learning algorithms and methods by participating in real research projects and solving practical problems that are important for various industries.

4) Skills in:

- Conducting independent scientific research and developing machine learning algorithms and models aimed at solving applied tasks.

- Applying effective teaching methods and interactive approaches that facilitate knowledge transfer and motivation to learn in machine learning and data science.

- Using modern information technologies, multi-threaded and multi-processor programming methods for implementing projects involving large data processing.

- Public speaking to convey scientific knowledge and research results to both academic and professional audiences, including at international scientific events.

- Professional communication and intercultural interaction necessary for working in international research projects and educational initiatives.

5) Competence in:

- Methods of organizing and conducting scientific research aimed at the practical application of machine learning and data analysis technologies in various industries.

- Modern educational technologies, including project-based learning and interdisciplinary research, which are applicable to tasks in artificial intelligence and data science.

- Developing and implementing educational programs and scientific projects aimed at applying solutions in machine learning and data analysis to address significant social and industrial challenges.

- Maintaining the relevance of one's knowledge and skills, engaging in continuous professional development, and advancing qualifications, which is essential for success in the rapidly evolving industry.

Professional Competencies:

- Mastering and Applying Knowledge – Ability to independently acquire, systematize, and apply knowledge to solve current challenges in data science and machine learning. This includes developing innovative approaches, adapting them, and applying them in research and applied activities.

- Setting Research Goals – Ability to formulate research objectives, develop strategies to achieve them, define the sequence of tasks, and choose methods necessary to obtain valid results and scientific conclusions.

- Statistical and Probabilistic Analysis – Ability to use methods of statistics, probability theory, and machine learning algorithms for scientific research and practical tasks, including the processing, analysis, and interpretation of large data volumes.

- Technical Skills and Tools – Proficiency in selecting and using scientific equipment, software tools, and instruments for data analysis and modeling, enabling effective data processing and high-quality results.

- Critical Analysis and Publication Activity – Ability to critically evaluate the results of one's professional activities and present them in scientific papers, conferences, and other publications, contributing to the development of science.

- Scientific and Technical Documentation – Ability to prepare and format scientific and technical documentation, including reports, reviews, and presentations to demonstrate research results, making them accessible to the scientific and professional community.

- Leadership and Inclusivity – Willingness to effectively manage a team, create a productive and inclusive work environment, and foster collaboration and respect for cultural and social differences.

- Professional Communication in a Foreign Language – Proficiency in a foreign language for presenting research results and participating in international projects, facilitating global knowledge exchange.

General Human and Socio-Ethical Competencies:

- Abstract Thinking and Critical Analysis – Ability to engage in abstract thinking, analysis, and synthesis, enabling the resolution of complex problems in data science and machine learning.

- Philosophical and Ethical Awareness – Applying philosophical knowledge to form a worldview and a critical approach to technology development, promoting conscious and ethical use of artificial intelligence.

- Socio-Cultural Understanding – Knowledge of social, cultural, and historical factors affecting societal processes, and the ability to analyze their impact on the use and implementation of technology in society.

- Economic Knowledge – Understanding of basic economics to assess the efficiency of projects and predict their consequences, including risk and cost evaluation in data analysis.

- Legal Awareness – Ability to consider legal aspects and intellectual property issues in professional activities, ensuring data protection and respect for the rights of all process participants.

- Self-Development and Creativity – Ability to develop one's professional skills and creative potential, supporting continuous learning and professional growth.

- Physical Culture and Stress Management – Applying physical activities to maintain health and manage stress in the context of intensive scientific work.

- First Aid and Safety – Proficiency in first aid and emergency response actions, ensuring safety and resilience in the workplace.

Special and Managerial Competencies:

- Ecological Analysis and Risk Management – Ability to analyze human impact on the environment, predict and assess potential risks arising when working with large data sets and developing algorithms.

- Designing Secure Systems – Ability to model processes and design systems with security requirements in mind, managing risks using analysis and forecasting methods.

- Risk Management Tools – Proficiency in modern software tools for modeling, evaluating, and managing risks, including big data analysis, ensuring reliability and stability in professional activities.

- Effectiveness and Compliance with Standards – Ability to assess and ensure compliance with legal and environmental standards, creating sustainable solutions in data processing and storage.

- Legal Competence – Knowledge of legal and regulatory requirements and the ability to apply them in system design, including evaluating social and environmental risks, ensuring the ethical and responsible use of technologies.

4.2 Requirements for the Doctoral Research Work

A doctoral student in the «Machine Learning & Data Science» program must possess the following competencies and skills to effectively conduct scientific research:

- Integration of Fundamental and Applied Knowledge – The ability to solve professional problems at the intersection of fundamental and technical sciences, combining theoretical knowledge with its practical application in software development and data science. This enables the creation of solutions that meet both modern scientific requirements and the practical needs of various industries.

- Independence in Scientific Research – The ability to conduct independent scientific research, analyze and interpret experimental data. The doctoral student must be able to formulate hypotheses, develop methodologies, conduct analysis, and present research results in accordance with international scientific standards, thereby contributing to the development of new knowledge and technologies.

- Modeling and Forecasting – The ability to develop and study models of the objects being researched, which enables performing complex analyses, simulations, and forecasts to solve practical problems. This skill allows doctoral students to conduct in-depth research and create models applicable to a wide range of data, from trend forecasting to process optimization.

- Proficiency in Modern Information Technologies – The ability to effectively use advanced information technologies for scientific research, including specialized software, databases, and tools for processing and analyzing large volumes of data. These skills allow doctoral students to identify significant dependencies in data and develop unique solutions for tasks in the fields of machine learning and data science.

These competencies allow doctoral students to make significant contributions to the development of the field of machine learning and data science, shaping them as researchers capable of finding innovative approaches to solving scientific and practical problems.

4.3 Requirements for the Organization of Practices

The educational program for doctoral students includes pedagogical and research practices aimed at developing and strengthening professional and scientific competencies:

– Pedagogical Practice – This practice is aimed at mastering teaching methods and skills in organizing the educational process. During this practice, doctoral students participate in educational activities at the undergraduate and master's levels, gaining experience in student interaction, developing mentoring skills, and strengthening their professional communication abilities. This experience helps doctoral students develop effective teaching skills and prepares them for work in educational institutions.

– Research Practice – This practice is organized to introduce doctoral students to modern scientific research methods, including data collection, processing, and analysis. It helps develop their research skills by allowing them to apply theoretical knowledge in practice and improve their competencies in scientific research. During the research practice, doctoral students learn methods for working with large data sets and apply advanced technologies to achieve high-quality scientific results.

These practices ensure comprehensive development of doctoral students, helping them acquire the necessary teaching and research skills, preparing them for successful professional activities and achieving significant results in the scientific field.

5 Requirements for completing studies and obtaining a diploma

To successfully complete the doctoral program in the field 8D06102 – «Machine Learning & Data Science», the doctoral student must accumulate at least 180 credits, distributed as follows:

- **45 credits** for theoretical training, covering disciplines necessary for developing deep knowledge in machine learning and data analysis.

- **123 credits** for pedagogical and research practices, aimed at developing both professional and research competencies.

- **12 credits** for the final assessment, which includes the defense of the dissertation research and confirms the doctoral student's qualification in the field of research and development.

Awarded Degree

Graduates of the program are awarded the academic degree of «Doctor of Philosophy (PhD)» in the field of 8D06102 – «Machine Learning & Data Science».

Graduate Competencies

Upon completing the doctoral program, the graduate should possess the following competencies:

General Professional Competencies:

- The ability to independently acquire, comprehend, and apply new knowledge and skills, fostering an innovative approach in professional activities.

- Skills in setting research goals, determining their sequence, and selecting methods to solve relevant tasks.

- Ability to apply knowledge from fundamental and applied disciplines that correspond to the program profile.

- The ability to critically analyze, present, and defend the results of their professional activities.

- Skills in preparing technical documentation, writing reports, reviews, presentations, and scientific publications.

- Readiness to lead a team, taking into account socio-cultural differences and principles of tolerance.

- Proficiency in professional communication in a foreign language.

Professional Competencies by Activity Types:

Engineering, Design, and Research Activities:

- Ability to perform highly complex technical developments in the field of machine learning.

- Ability to design and analyze algorithms and models for various machine learning applications.

- Skills in applying data analysis methods and processing large volumes of information.

- Ability to use modern tools and programming languages for model development and optimization.

- Ability to design and develop machine learning systems to solve specialized tasks.

Production and Technological Activities:

- Proficiency in independently performing project and scientific-production work in machine learning.

- Ability to use methods for operating modern equipment and developing tools.

- Ability to apply data processing methods to solve applied production tasks.

- Skills in testing and validating machine learning models.

- Ability to manage the lifecycle of models, including data preparation, training, and deployment stages.

Organizational and Managerial Activities:

- Readiness to organize and manage research and production projects in the field of machine learning.

- Ability to apply regulatory and methodological documents when planning model development.

- Ability to interact with clients and partners, using standards for project and process documentation.

- Ability to assess and choose approaches to model development and integration, considering risk management.

- Proficiency in decision-making and data analysis methods for optimizing model development.

Scientific and Pedagogical Activities:

- Skills in conducting seminars, laboratory, and practical classes in machine learning.

- Participation in the development of educational and methodological materials and documentation.

- Ability to supervise the scientific and academic activities of students in the field of machine learning.

Program Learning Outcomes

The doctoral program provides graduates with the necessary cultural and professional competencies, preparing them for successful professional and scientific research activities in the fields of machine learning and data analysis.

CONCLUSION

The doctoral program in the field 8D06102 – «Machine Learning & Data Science» is aimed at training highly qualified scientists and researchers capable of solving complex scientific problems in the areas of machine learning and data science. Throughout the program, doctoral students acquire the skills necessary to conduct independent research, critically analyze data, and develop advanced models and algorithms applicable across various fields.

Graduates of the program possess competencies that include a deep understanding of theoretical foundations, practical data processing skills, as well as managerial and pedagogical abilities. This enables them not only to thrive in scientific activities but also to participate in educational processes, making them highly sought-after professionals on the international stage.

The program is focused on developing research potential, innovative thinking, and leadership qualities, contributing to the formation of competent and responsible specialists ready for independent scientific and professional activities. Thus, graduates of the 8D06102 – «Machine Learning & Data Science» program become not only experts but leaders capable of making significant contributions to the advancement of science and technology on a global scale.

Head of the Department of   _____   **Abdoldina F.N.**
«Software Engineering»