



## MODERN TRENDS IN INTELLIGENT DATA ANALYSIS

G.N.Aimbetova<sup>1</sup>

Senior lecturer

N.D.Bekniyazova<sup>2</sup>

Assistant

S.B.Izemetov<sup>3</sup>

Assistant

Department of Nukus State Technical University

<https://doi.org/10.5281/zenodo.15551017>

### ARTICLE INFO

Received: 24<sup>th</sup> May 2025

Accepted: 29<sup>th</sup> May 2025

Online: 30<sup>th</sup> May 2025

### KEYWORDS

Data Mining, machine learning, AutoML, in-depth learning, explainable AI, graphical analysis, streaming data processing, federated learning.

### ABSTRACT

*This article explores the current trends in modern data mining, emphasizing its growing automation, complexity, and user-friendliness. Key developments such as Automated Machine Learning (AutoML), Deep Learning, Explainable AI (XAI), Graph-Based Analysis, Stream Data Processing, Federated Learning, and Quantum Computing are reshaping how data is collected, processed, and interpreted. These technologies enable more efficient, transparent, and privacy-preserving data analysis across various domains, including healthcare, finance, and social networks.*

### Introduction

Data Mining is crucial for working with large volumes of data and artificial intelligence. This field is rapidly developing alongside fields such as artificial intelligence (AI), machine learning (ML), and cloud computing, and methods for extracting knowledge from this data in these fields are developing year by year. What innovations do Data Mining technologies have today? In this article, we will consider the most relevant trends in data analysis.

#### 1. Automated Machine Learning (AutoML)

The automated process of machine learning is among the main trends, allowing AutoML (Automated Machine Learning) to build a model, optimize parameters, and automatically select algorithms. AutoML can provide great opportunities for simple analysts and small teams analyzing data in companies.

Well-known tools include Google AutoML, H2O.AI, and TPOT (Tree-based Pipeline Optimization Tool).

Google AutoML is the process of creating a dynamic combination of various methods to form an entire conveyor of easy-to-use machine learning. In other words, it's Google's new product that combines various ready-made machine learning devices that have been trained to a certain degree. It is based on the Neural Architecture Search (NAS) algorithm for neural architecture search. This algorithm not only recognizes images, but also searches for and uses the most suitable neural network for solving a problem in cloud technology.

#### 2. Deep Learning of Data Mining

Deep Neural Networks, especially transformers (e.g., BERT, GPT), are crucial for analyzing images, text, and sounds. Currently, transformer models (BERT, GPT) are showing successful results.



When it comes to deep learning, it is also worth mentioning artificial intelligence (AI) and its closely related Machine Learning (ML). If machine learning is a way to achieve artificial intelligence, deep learning is a technology to achieve machine learning, that is, deep learning is a machine learning algorithm that meets the characteristics. The concept of its study arises from the analysis of neural networks.

Deep learning is used in natural language processing or text comprehension (NLP), computer vision or image recognition (Computer Vision), and predictive analytics (Predictive Analytics).

### **3. Interpreting and explaining artificial intelligence (Explainable AI, XAI)**

Explanatory AI (XAI) is a part of artificial intelligence research that strives to create systems and models that explain their actions and make decisions in a way that people understand. One of the main problems in modern artificial intelligence training is that "black box" systems can provide accurate answers and perform complex tasks, but it is often difficult to understand how they achieved such results. In response to this, explanatory artificial intelligence (XAI) was created, the purpose of which is to make decision-making in AI systems convenient and easy.

XAI (Explainable AI) is designed to create interpretable models and is used to understand model solutions in the systems of medicine, finance, and jurisprudence.

Decisions can be made using SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model-Agnostic Explanations), and Decision Trees.

### **4. Graphical Mining**

Social networks, banking transaction systems, and biological data are usually presented as graphs. Modern methods of graphical analysis help to identify society, influential agents, and anomalies.

Graph mining (graphical analysis) is a set of methods and algorithms for extracting useful data from general graphical structures.

The main algorithms and methods include Frequent subgraph mining, Graph clustering, Graph classification, Link prediction, and Community detection.

Graphical analysis is mainly used in analyzing social networks (LinkedIn, Facebook), for example, in analyzing the spread of trends.

Graph is also used in bioinformatics to study protein interactions, fraud detection, and analysis of financial fraud networks.

Graph mining is one of the most promising areas of modern data science and is a powerful tool for extracting valuable information from large volumes of data.

### **5. Stream Mining (Data Processing)**

With the year-on-year development of IoT and online transactions, the need for real-time data analysis is growing. Stream mining technologies (Apache Kafka, Flink) serve this purpose.

Stream Mining (Data Processing) is the process of extracting useful data from a large and continuous stream of data. This method is used to analyze frequently changing data over time. As an example, we can cite data from sensors (meteorology, GPS, IoT tools), online transactions (banking operations, e-commerce), social media streams (Twitter, Facebook posts), and traffic analysis in computer networks. The main algorithms include:



1. Sliding Window model - looks at only a portion of the newly received data.
2. Sampling - stores only a short portion of the data and performs rapid analysis.
3. Sketching (Brief Writing) - works quickly by shortening data (for example, Bloom Filter).
4. Online clustering groups data over time.

Widely used areas of streaming mining are:

- ✓ Fraud Detection (Finding Fraud) - quickly detects events in banking operations;
- ✓ Finding trends - finds trending topics on social media;
- ✓ Sensory analysis is used in agronomy, medicine, and the automotive industry.
- ✓ Network security is the prevention of cyberattacks.

Stream Mining is a crucial technology for analyzing large and rapidly changing data. Working with classical databases is not recommended, as they cannot process large and dynamic data. The main goal of Stream Mining is quick, unlimited, and effective analysis.

## 6. Federated Learning

Instead of collecting data on a centralized server, federated learning (Federated Learning) allows you to train models on decentralized devices (smartphones, IoT devices) while maintaining confidentiality.

Federated Learning (FL) is a method of using multiple tools (telephones, computers, IoT devices) to collaboratively study a model without sharing their data.

The main principle of federative learning is:

- ✓ Local learning - each tool learns its own information and updates the model.
- ✓ Combining updated data - updated data from all sources will be combined in one center.
- ✓ Distribution of the new model - the updated model will be resubmitted to all vehicles.
- ✓ The process is repeated - the model continues to improve.

The advantages of federative learning include:

1. GDPR (General Data Protection Regulation) - the general rule for data storage and compliance with other rules;
2. Internet traffic can be saved;
3. Data protection is much higher

The shortcomings of Federated Learning can also be noted. For example, data protection, meaning user data is not sent to a centralized server.

Reducing traffic, in which big data is not exchanged, only an updated version of the model is sent. Learning is carried out in real time, meaning tools learn their information over time.

As areas of application, we can cite smart phones (in photo recognition), in medicine (diagnosis based on patient data), in the banking system (training a fraud detection model).

There are several types of federated learning, including horizontal, vertical, and federated transfer learning, which involves working with various structured data.

## 7. Using Quantum Computing

Although the technology is still in its early stages, quantum computing holds great promise for the future of data mining. Quantum algorithms, such as quantum support vector machines (quantum SVM), have the potential to significantly accelerate the solution of complex optimization problems that are difficult or time-consuming for classical computers. With the ability to process and analyze massive datasets in parallel using quantum bits



(qubits), quantum computing can enhance the efficiency and scalability of data mining tasks. As quantum hardware and software continue to evolve, their integration into data science workflows may open up entirely new horizons for intelligent data analysis, especially in fields like cryptography, drug discovery, and financial modeling.

## Conclusion

Modern data mining is becoming increasingly automated, complex, and user-oriented. Trends such as AutoML, deep learning, and federated learning are creating new opportunities for businesses, science, and society to address current challenges. In the future, data mining, artificial intelligence, and quantum technologies are expected to become even more closely interconnected.

The discussed trends—Automated Machine Learning (AutoML), Deep Learning, Explainable AI (XAI), Graph-Based Analysis, Stream Data Processing, Federated Learning, and the emerging role of Quantum Computing—highlight a significant shift in how data is collected, processed, and interpreted.

AutoML enables even non-experts to develop powerful models with minimal manual intervention. Deep learning, particularly with transformer-based architectures, has revolutionized how we process natural language, images, and audio. Explainable AI ensures transparency and trust in model decisions, especially in critical fields such as healthcare and finance.

Graph-based analysis offers a powerful approach to modeling complex relationships like those found in social networks and financial systems. Stream mining addresses the challenge of real-time data generated by sensors, IoT devices, and online transactions, enabling faster and more accurate decision-making. Federated learning ensures data privacy and regulatory compliance by training models across decentralized devices. Meanwhile, quantum computing—though still emerging—holds the potential to dramatically accelerate data mining tasks in the near future.

This article provides an overview of the most relevant areas in modern data mining. The insights presented here will be valuable for analysts, IT professionals, and anyone interested in data technologies.

## References:

1. Radford, A. (2021). "Learning Transferable Visual Models From Natural Language Supervision."
2. Brown, T.(2020). "Language Models are Few-Shot Learners."
3. Luckin, R. (2018). "Machine Learning and Human Intelligence: The Future of Education."
4. Aimbetoba G.N., Axunov F., Li A. Features of developing a crm system for medical centers using mysql, php and javascript technologies// Central asian journal of mathematical theory and computer sciences 2021год 31.01.