Proceedings of Machine Learning Research vol 278:1–13, 2025 2025 2nd International Conference on Machine Learning and Intelligent Computing

Research on Data Mining Techniques Based on DeepSeek-R1

Lei Cai* CLEI1005@163.COM
The 714th Research Institute of China State Shipbuilding Corporation Limited

Fan Yin 11899342@QQ.COM
The 714th Research Institute of China State Shipbuilding Corporation Limited

Xianbo Meng 48198320@QQ.COM
The 714th Research Institute of China State Shipbuilding Corporation Limited
**Corresponding author*

Editors: Nianyin Zeng, Ram Bilas Pachori and Dongshu Wang

Abstract

With the rapid development of artificial intelligence technology, data mining, as one of its important application fields, is facing new opportunities and challenges. DeepSeek-R1, as an advanced pre-trained model, provides new technical means for data mining with its powerful feature extraction capabilities and efficient inference performance. This paper systematically investigates data mining techniques based on DeepSeek-R1, offering a comprehensive exploration of technical principles, application methods, and performance optimization. Experimental results demonstrate that DeepSeek-R1 exhibits significant performance advantages in data mining tasks, and corresponding optimization strategies are proposed. The research in this paper not only enriches the theoretical system of data mining technology but also provides valuable references for practical applications.
Keywords: Artificial Intelligence; Large Models; Data Mining

1. Introduction

Data mining refers to the process of extracting valuable information from large volumes of data and is widely applied in fields such as business, finance, healthcare, and scientific research (Hand, 2007). With the explosive growth of data volume and the diversification of data types, traditional data mining methods face numerous challenges, such as difficulties in feature extraction, handling high-dimensional data, and insufficient model generalization capabilities. These limitations have spurred the need for more advanced techniques (Han et al., 2011).

In recent years, the rapid development of artificial intelligence technologies, particularly deep learning, has brought new solutions to data mining (Goodfellow et al., 2016). For instance, (Zhang et al., 2024c) integrated the Gemini series and GPT-4 series models into medical data analysis, evaluating their efficiency and accuracy in interpreting and utilizing visual information, which highlights the potential of multi-modal approaches (Zhang et al., 2024b). Meanwhile, Zhang et al. (2024b) focused on mining structured knowledge from textual data by fine-tuning large language models (LLMs) using human-annotated training data, thereby achieving weakly supervised information extraction to extract entity and relational structures from text (Zhang et al., 2024c), demonstrating the adaptability of LLMs in diverse tasks. Zhang et al. (2024a) applied GPT-3.5-turbo to the task of knowledge extraction from chemical texts, which is widely regarded as a highly challenging endeavor. The model achieved accuracy levels ranging from 69% to 95% across tasks such as compound entity recognition and reaction role tagging.

DeepSeek-R1, a pre-trained model based on deep learning, has become a new favorite in the field of data mining due to its powerful feature extraction capabilities, robustness in handling complex data patterns, and efficient inference performance (Chen and Zhang). This paper aims to explore the application of DeepSeek-R1 in data mining and its performance optimization strategies, with the goal of providing references for research and practice in related fields. By analyzing its strengths and addressing potential limitations, this study seeks to pave the way for more effective data-driven decision-making processes.

2. DeepSeek-R1 Technical Principles

2.1. Key Technologies of DeepSeek-R1

DeepSeek-R1 achieves breakthroughs in reasoning efficiency, training stability, and adaptability to multiple scenarios through technologies such as MLA+MoE architecture, GRPO reinforcement learning, and dynamic reward design. Its technical approach provides a new paradigm for the lightweight deployment and specialized training of large-scale language models (Rafailov et al., 2023).

In terms of DeepSeek-R1's architecture, it adopts the Mixture of Experts (MoE) architecture. MoE decomposes complex problems into multiple subtasks, which are handled by different expert networks. These experts are small neural networks trained for specific domains or tasks, such as grammar, factual knowledge, or creative text generation. MoE only activates the experts relevant to the current task, which significantly reduces computational costs while improving efficiency. The MoE architecture can be formally represented by the following formula (Chen and Zhang):

$$y = \sum_{i=1}^{N} G_i(x) E_i(x) \tag{1}$$

Where,

- N is the number of experts.
- $G_i(x)$ is the weight assigned to the i-th expert by the gating network, typically satisfying $G_i(x)$.
- $E_i(x)$ is the computation result of the i-th expert.

Unlike traditional MoE, DeepSeek-R1 divides experts into finer granularities, reducing the parameter size of each expert while increasing the number of experts. This ensures that the total parameter size and the number of activated parameters in the MoE module remain unchanged, while also allowing for more flexible combinations of multiple experts. Additionally, DeepSeek-R1 distinguishes activated experts into **Shared Experts** and **Routed Experts**, which have significant differences in data processing workflows. For Routed Experts, input data first passes through a routing module, which selects the most suitable expert for computation based on the features of the input data. Finally, the computation results of the Routed Experts and Shared Experts are combined to form the final output of the MoE module. This design enhances the model's generalization capability and adaptability. To address the common issue of load imbalance in MoE, a novel load balancing strategy is proposed. A learnable bias term is introduced into the Gate module used for expert selection. When calculating routing scores, this bias term is dynamically added to the score of each Routed Expert, allowing for dynamic adjustment of routing tendencies without additional overhead (Vaswani et al., 2017).

From the training perspective, DeepSeek-R1 adopts the **Group Relative Policy Optimization (GRPO)** scheme, which can be considered a computationally efficient version of PPO (Proximal Policy Optimization). GRPO maintains performance while reducing computational resource consumption. PPO is widely regarded as a benchmark algorithm in the field of reinforcement learning, utilizing four models: the Policy model (also known as the Actor) and the Value model (also known as the Critic). In the context of natural language processing (NLP) generation models, the model being trained acts as the actor, and its performance corresponds to the generated responses. The PPO algorithm can be formally represented by the following formula (Gu et al., 2022):

$$L^{clip}(\theta) = E_t \left[\min \left(r_t(\theta) A_t, \mathrm{clip}\left(r_t(\theta), 1-\varepsilon, 1+\varepsilon\right) A_t\right)\right] \tag{2}$$

Where:

- θ is the update policy, often referring to the parameters of the neural network.
- $r_t(\theta)$ is the policy update ratio.
- A_t is the advantage function, representing the relative advantage of taking an action in a given state.
- ε is a hyperparameter used to limit the magnitude of policy updates, typically taking a value between 0.1 and 0.3.
- $\mathrm{clip}\left(r_t(\theta), 1-\varepsilon, 1+\varepsilon\right)$ restricts the policy update ratio to the range $1-\varepsilon, 1+\varepsilon$.

The GRPO algorithm differs from the PPO algorithm in that it estimates the advantage among a group of data rather than between two data points. This allows for faster policy updates and more accurate advantage estimation.

DeepSeek-R1 extensively applies reinforcement learning in model training, optimizing the model's decision-making capabilities through a trial-and-error mechanism and environmental feedback. Additionally, a rule-based reward system has been developed to guide the model's learning process, enhancing training efficiency and logical reasoning abilities (Rafailov et al., 2023).

In terms of data processing, DeepSeek-R1 can integrate multimodal data. In scenarios such as object detection in autonomous driving within the computer vision field and lesion identification in medical image analysis, it can provide more accurate diagnostic results. The model employs knowledge distillation techniques to compress the capabilities of larger models into smaller-scale models. This technology enables DeepSeek-R1 to remain competitive in hardware-constrained environments, with some model parameters being as few as 1.5 billion while still capable of performing complex tasks (Chen and Zhang).

These technical features bring benefits such as improved computational efficiency, faster inference speeds, reduced costs, and enhanced cost-performance advantages to DeepSeek-R1.

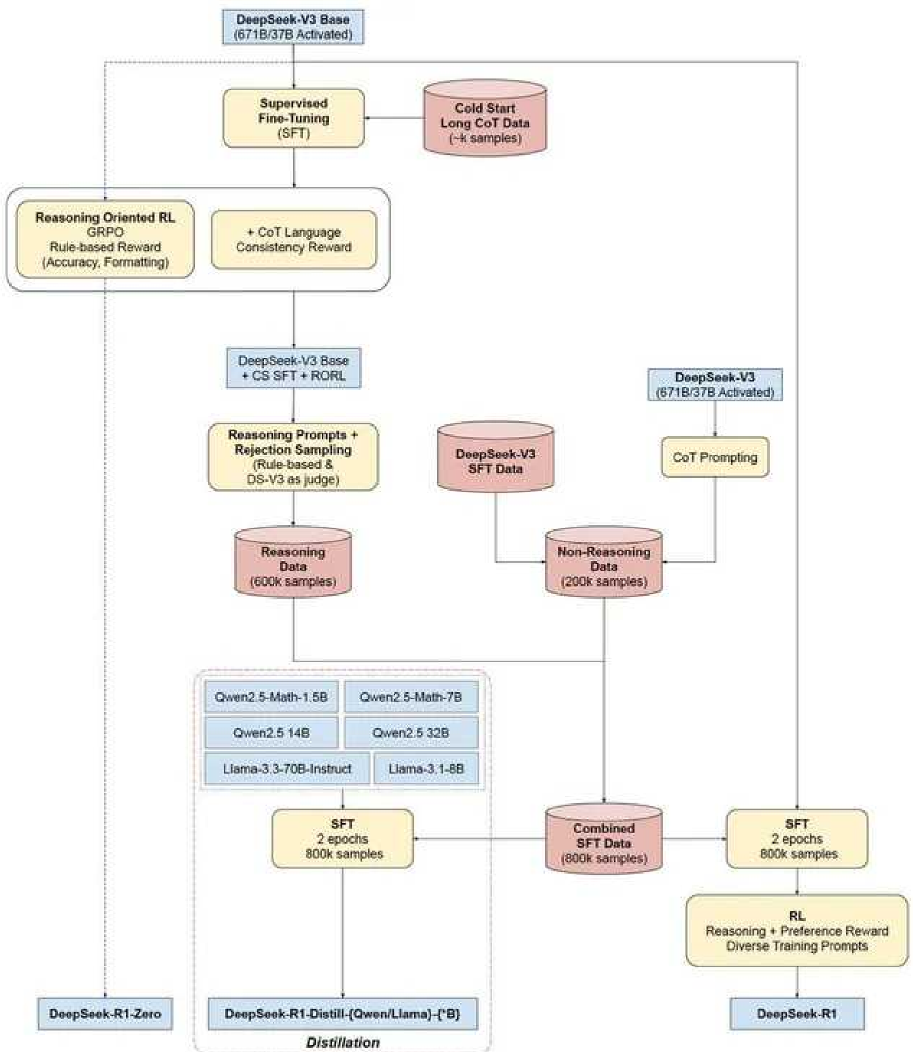

Figure 1: DeepSeek R1 Training Flowchart.

2.2. Training Process of DeepSeek-R1

DeepSeek-R1 is not trained from scratch but is based on the DeepSeek-V3 model, As shown in Figure 1, Its training process is as follows (Yu et al., 2025):

Step 1: Base Model Selection

The replication of DeepSeek-R1 requires a base model and a model capable of generating long-chain CoT (Chain-of-Thought) data. The current plan is to use DeepSeek-V3 (DS V3) and DeepSeek-R1-0 (DS R1-0) as the base models for training.

Step 2: First SFT (Supervised Fine-Tuning) (Cold Start Phase)

Data Preparation:

- Data Generation:The long-chain CoT data for SFT is generated by R1-Zero and validated by humans.

- The human validation process includes (Murphy, 2012):

(1) Correctness of Special Tokens: Ensuring that special tokens in the data comply with predefined standards.

(2) Coverage of Knowledge: Ensuring the data covers a sufficiently broad range of knowledge domains.

(3) Accuracy of Knowledge: Verifying the correctness of reasoning and facts.

Training Process Components:

- Training Algorithm: The classic "autoregressive loss" is used.

- Training Objective: Address the poor readability and mixed language output issues of R1-Zero.

- Training Results: Through carefully designed cold-start data with prior human patterns, the model at this stage exhibits better performance than R1-Zero.

Step 3: First RL (Reinforcement Learning)

Data Preparation:

A dataset not publicly available from DeepSeek is used here. This dataset was used to train DeepSeek-V3 (DS V3) into DeepSeek-R1-0 (Reasoning data). Since this data is not accessible, high-quality data can be obtained through "CoT distillation" from the existing DeepSeek-R1 model, followed by repeated rejection sampling.

Training Process Components:

- Training Algorithm: "GRPO (Group Relative Policy Optimization)" is employed, combined with a rule-based reward mechanism.

- Training Objectives include:

(1) Enable the model to express reasoning processes more clearly.

(2) Generate enhanced long-chain CoT outputs that conform to "+COT language."

(3) Improve the granularity of reasoning steps, the ability to express logical relationships, and the optimization of language style.

- Training Results:The generated model enhances the logical consistency and expressive quality of CoT while more accurately capturing long-chain reasoning processes.

Step 4: Generation of Non-Reasoning Samples

Data Preparation:

- Data reuse:Reuse DeepSeek-V3 SFT data.

- Data Generation: Use DS V3 (671B or 37B) to generate chain-of-thought prompt data.

- Process non-reasoning task data specifically (simple queries do not generate CoT, while complex tasks generate chain-of-thought).

Training Process Components:

-Training Algorithm: Use the DeepSeek-V3 process for data generation.

-Training Objective: Generate "200,000 high-quality non-reasoning samples."

Step 5: Generation of Reasoning Samples

Data Preparation:

Use the output data from the SFT-trained model in Step 2, along with additional data, to create a training set.

Cleaning Algorithm:

Reasoning prompts + rejection sampling:

- Address the issue of insufficient reasoning capabilities in the model.

- Handle reasoning-intensive tasks to improve the clarity of the model's reasoning.

- Mitigate language confusion and eliminate redundant information.

- Use DeepSeek-V3 as a judge for automated evaluation to filter high-quality data.

Data Objective:

- Generate many CoT samples through rejection sampling (Yang et al., 2025)..
- Resolve issues such as reasoning confusion, lengthy paragraphs, and redundant code.

Generation Results:

Generate “600,000 high-quality reasoning samples.”

Step 6: Data Integration

Combine “600,000 reasoning samples” with “200,000 non-reasoning samples” to form “800,000 high-quality sample data.”

Step 7: Model Distillation

Training Objective:

Perform “one round of SFT training” using all the integrated data as SFT data to fine-tune a specified small model, such as Qwen 7B.

Training Results:

- Fully utilize enhanced CoT data for fine-tuning.
- Improve the reasoning capabilities of the small model, giving it reasoning abilities similar to the R1 model.

3. Implementation of Data Mining Technology based on DeepSeek-R1

The implementation of data mining technology based on DeepSeek-R1 primarily includes the data mining process and the model training process. The data mining process involves data processing, algorithm optimization, data mining, and graphical representation. The model process is based on distilling DeepSeek using the data to form a domain-specific model suitable for the current data mining scenario. Leveraging the model's capabilities, it enhances the data cleaning, data integration, data fusion, and algorithm components within the data mining process. This approach improves data processing capabilities and enhances algorithmic performance. As shown in Figure 2.

3.1. Application in the Data Preprocessing Phase

Throughout the entire data mining process, data preprocessing can take up to 60% of the time. However, data that has undergone preprocessing not only saves a significant amount of space and time but also yields mining results that are more effective for decision-making and prediction purposes (Zhang et al., 2024b).

Data preprocessing is divided into several stages: data collection, data cleaning, data integration, and data fusion, as illustrated in the following diagram (Rastogi and Bansal, 2023):

As shown in Figure 3, In the realm of data cleaning, DeepSeek R1 proves to be highly effective in removing noise from text data, such as special characters, punctuation marks, and HTML tags. When processing text scraped from web pages, it precisely eliminates these messy tags and superfluous spaces, rendering the text clean and standardized, thereby laying a solid foundation for subsequent analysis and processing. In terms of format standardization, it can convert data of various formats into a unified format, such as transforming different date representations into the “YYYY - MM - DD” format, facilitating data integration and further analysis (Rastogi and Bansal, 2023).

Data transformation is a crucial step in data preprocessing, and DeepSeek - R1 excels at converting raw text into vector representations suitable for model processing. Utilizing technologies like the Word2Vec model, it transforms words in the text into vectors, converting textual data into

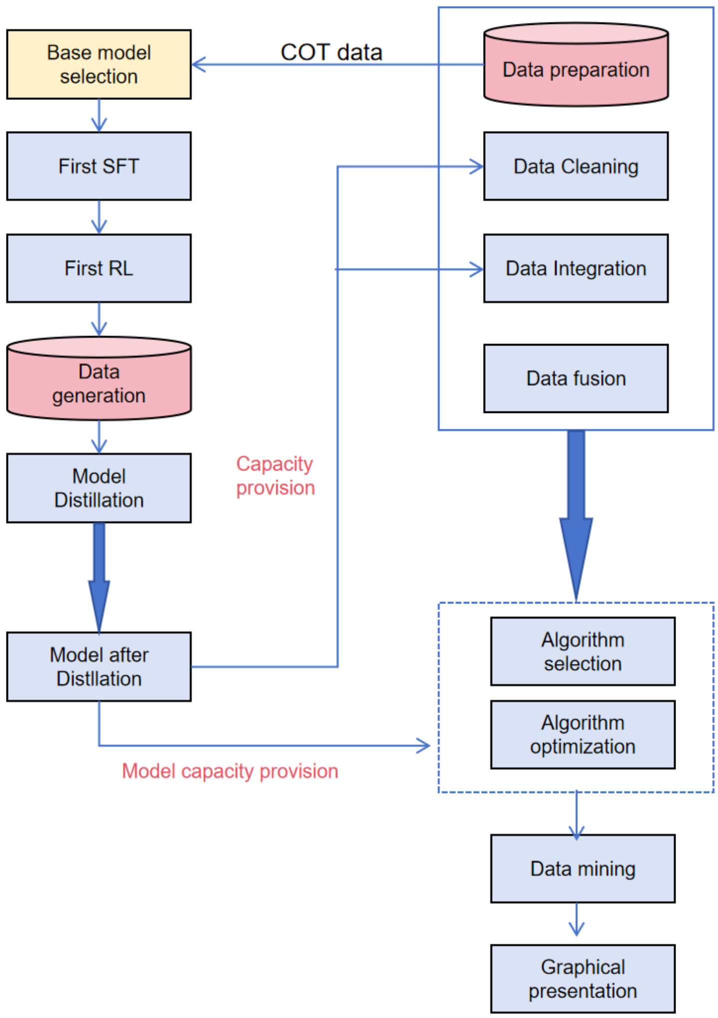

Figure 2: Distillation Model Flowchart.

numerical data to meet the input requirements of models. Additionally, based on model needs, it can also perform data type conversions, such as changing character data into numerical data or discretizing numerical data, making the data more compatible for model training and prediction.

During the data integration phase, it is essential to mine key information from the data through feature extraction and organically fuse the data. DeepSeek - R1 showcases its strengths in feature extraction. In keyword extraction, leveraging its natural language processing capabilities, DeepSeek - R1 employs methods like word frequency statistics and TF-IDF to accurately extract keywords from text data. When dealing with news articles, by extracting core keywords, it can quickly grasp the article's theme and key points, facilitating rapid information screening and analysis. Semantic understanding is another vital function, enabling deep semantic analysis of text data to comprehend the meaning and contextual information of the text. In sentiment analysis tasks, through semantic modeling, it can accurately determine whether the sentiment expressed in the text is positive, negative, or neutral. Beyond text feature extraction, DeepSeek - R1 can also extract features from

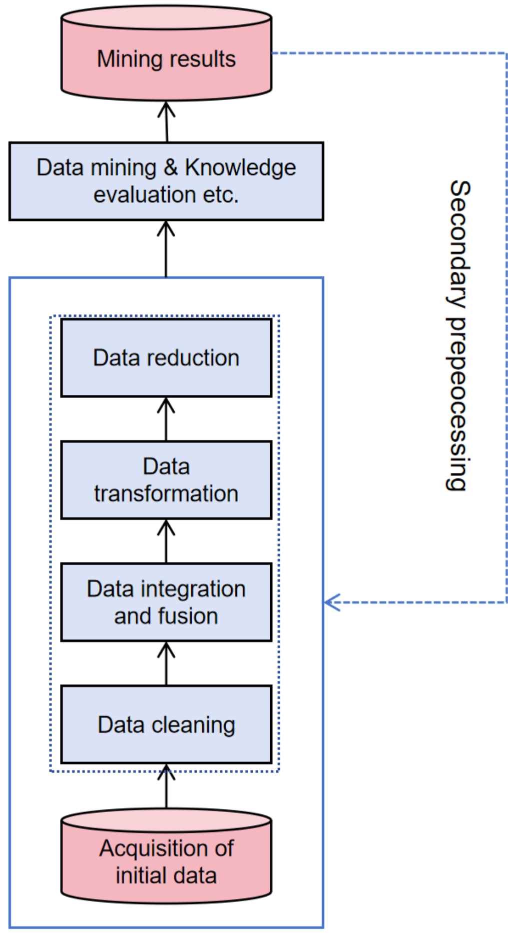

Figure 3: Data Preprocessing Steps.

other types of data. For image data, it can extract color features, texture features, etc.; for audio data, it can extract frequency features, duration features, etc. These multi-type feature extraction capabilities ensure that data from various sources can be effectively preprocessed, enhancing data usability and model performance (Li and Zhang, 2024).

DeepSeek - R1, through data cleaning, removes noise and standardizes formats; through data transformation, it adapts data to model requirements; and through feature extraction, it mines key information from the data. These functions work in concert to provide high-quality, usable data for data analysis and model training.

3.2. Selection and Optimization of Data Mining Algorithms

After completing data preprocessing, the next step is to proceed with data mining. Selecting an appropriate data mining algorithm is crucial, as it enables the rapid and accurate identification of valuable information from vast and complex datasets from a precision analysis perspective. In terms of predictive capabilities, algorithms can construct models based on historical data to forecast future trends. Time-series algorithms, for instance, can predict stock price movements or weather changes, providing a scientific basis for decision-making. In the realm of data classification and clustering, classification algorithms can accurately categorize data, aiding users in efficient data management and comprehension; clustering algorithms, on the other hand, can group similar data together, uncovering the intrinsic structure of the data. Thus, data mining algorithms are key to unlocking the value of data and advancing decision-making scientificity across various fields (Wang and Zhao, 2024).

There are a wide variety of data mining algorithms, such as classification algorithms, clustering algorithms, and association rule mining algorithms, each with its applicable scenarios and limitations. DeepSeek - R1 assists in selecting the appropriate algorithm. Below is an example illustrating how DeepSeek_R1 aids in algorithm selection.

Understanding the characteristics of the data is fundamental to choosing the right algorithm. DeepSeek - R1 can conduct a comprehensive analysis of the dataset, including the data types (such as numerical, textual, image-based, etc.), data distribution (such as normal distribution, skewed distribution, etc.), and data scale (large datasets, small datasets). For example, when faced with a dataset containing many customer transaction records, DeepSeek - R1 can analyze that the dataset includes numerical features like transaction amounts and quantities, and textual features like product descriptions. Based on these characteristics, it be can determine that if the goal is to predict customer purchasing behavior, a classification algorithm might be suitable; if the aim is to discover potential patterns among customer groups, a clustering algorithm might be more appropriate.

Business requirements are also a critical factor in algorithm selection. DeepSeek - R1 can understand the user's business objectives and translate them into specific data mining tasks. For instance, in the field of marketing, a user's business need might be to identify potential high-value customer groups. DeepSeek - R1 can recommend suitable clustering or classification algorithms to achieve this goal, taking into account factors such as the algorithm's interpretability, accuracy, and efficiency, and providing the user with the best algorithm selection advice after comprehensive evaluation.

DeepSeek - R1 can also refer to the experiences and best practices of similar past projects. It can search its vast knowledge base for cases similar to the current dataset and business requirements, and analyze the algorithms used in these cases and their outcomes, thereby offering more targeted algorithm selection references to the user.

The ten classic data mining algorithms include: C4.5, Support Vector Machines (SVM), Apriori algorithm, Expectation–Maximization (EM) algorithm, PageRank, Adaboost, k-Nearest Neighbor (KNN) classification algorithm, Decision Tree Model, and Naive Bayesian Model (NBC).

Below is an example of how DeepSeek-R1 can optimize the K-Means algorithm. The K-means algorithm is a commonly used clustering analysis method aimed at partitioning n data points into k clusters, such that each data point belongs to the cluster with the nearest mean (i.e., cluster center), thereby minimizing the sum of squared errors within the clusters. The basic steps of the K-means algorithm are as follows:

1. Initialization: Randomly select k data points as initial cluster centers.
2. Assignment step: Assign each data point to the nearest cluster center, forming k clusters.
3. Update step: Recalculate the center of each cluster as the mean of all points within the cluster.

4. Repeat steps 2 and 3: Until the cluster centers no longer change or a preset number of iterations is reached, at which point the algorithm converges.

$$J = \sum_{k=1}^{K} \sum_{i=1}^{n} \|x_i - \mu_k\|^2 \tag{3}$$

The performance of the K-means algorithm largely depends on the selection of initial centroids. Randomly choosing initial centroids may lead the algorithm to fall into local optima, affecting the clustering outcome. DeepSeek, with its powerful natural language processing and data analysis capabilities, can comprehensively understand and analyze datasets. By learning from data distribution, feature importance, and other information, the large model can recommend more suitable initial centroids. It first performs preliminary clustering or classification predictions on the data and selects representative samples from each category as initial centroids based on the prediction results, thereby improving the algorithm's convergence speed and clustering quality.

Before using the K-means algorithm, it is necessary to predefine the number of clusters, k. Determining the appropriate value of k is often challenging. An inappropriate choice of k may lead to unreasonable clustering results. DeepSeek can assist in determining the suitable number of clusters through semantic understanding and pattern recognition of the data. The large model can analyze the relationships between different features in the dataset and, considering business needs and data characteristics, suggest a reasonable range for k. Additionally, the large model can simulate clustering results under different k values to evaluate the quality of clustering, helping users select the optimal k value.

3.3. Domain Model Distillation Training

From the training process of DeepSeek-R1 outlined above, we can observe the entire distillation process of R1. Following this workflow, it is also possible to perform distillation training on a large model based on the data objects in data mining, resulting in a domain-specific model that enhances the model's accuracy in empowering key aspects of data mining.

1. Data Preparation

- Data Collection: Gather task-relevant data from various sources (such as databases, file systems, network interfaces, etc.). For example, if the task is image classification, a large number of images from different categories need to be collected.

- Data Preprocessing: Clean the collected data (handling missing values, outliers, etc.), transform it (e.g., normalization, standardization), and split it (dividing the data into training, validation, and test sets). For instance, in a house price prediction task, features like house area and number of rooms may need to be normalized, and the dataset should be split into proportions (e.g., 70% training set, 15% validation set, 15% test set). 2. Model Selection

- Task Matching: Choose an appropriate model based on the specific task type (classification, regression, clustering, etc.). For example, for text classification tasks, models like Naive Bayes, Support Vector Machines, or deep learning models such as Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN) and their variants (e.g., LSTM, GRU) can be selected.

- Model Complexity: Consider the data scale and feature count to select a model with suitable complexity. For smaller datasets, overly complex models may overfit, while for larger datasets, simpler models might not capture the information adequately.

3. Model Training

- Hyperparameter Setting: Different models have different hyperparameters, such as learning rate, number of iterations, regularization coefficients, etc. Methods like grid search or random search can be used to find the optimal combination of hyperparameters. For example, when training a neural network, hyperparameters like learning rate, batch size, and number of epochs need to be set.

- Training Process: Train the model using the training set, continuously adjusting the model's parameters through optimization algorithms (e.g., gradient descent, stochastic gradient descent) to minimize the loss function on the training set.

4. Model Validation

- Using Validation Set: During training, use the validation set to evaluate the model's performance and monitor for overfitting or underfitting. If the model performs well on the training set but poorly on the validation set, it may be overfitting; if it performs poorly on both, it might be underfitting.

- Model Adjustment: Adjust the model based on the validation set's evaluation results, such as tweaking hyperparameters or modifying the model's complexity.

5. Model Evaluation

- Selecting Evaluation Metrics: Choose appropriate evaluation metrics based on the task type. For classification tasks, common metrics include accuracy, precision, recall, and F1 score; for regression tasks, metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) are used.

- Using Test Set: Evaluate the final trained model using the test set to determine its true performance on unseen data.

4. Conclusions and Future Prospects

This paper systematically investigates data mining techniques based on DeepSeek-R1, providing a comprehensive exploration of its technical principles, application methods, and performance optimization. The research results demonstrate that DeepSeek-R1, with its advanced architectural design (such as MLA+MoE architecture, GRPO reinforcement learning, and dynamic reward design) and efficient training processes, exhibits significant performance advantages in data mining tasks. During the data preprocessing phase, it effectively removes noise, standardizes data formats, and extracts key information through feature extraction. In terms of data mining algorithm selection and optimization, DeepSeek-R1 can recommend suitable algorithms based on data characteristics and business requirements, and enhance model performance by optimizing algorithm parameters. Furthermore, through domain model distillation training, DeepSeek-R1 can further adapt to specific data mining scenarios, improving the model's inference capabilities and generalization ability.

Despite the notable progress made by DeepSeek-R1 in the field of data mining, there are still areas worthy of further research and improvement. First, as data scale and complexity continue to grow, enhancing the model's computational efficiency and inference speed remains a challenge. Future work could explore more efficient architectural designs and training strategies to meet the demands of processing larger-scale data. Additionally, model interpretability remains a critical issue. In practical applications, users often need to understand the decision-making process and rationale behind the model. Therefore, future research could focus on integrating technologies such as knowledge graphs and causal reasoning to improve model interpretability, making it more trustworthy in complex tasks.

In conclusion, DeepSeek-R1 brings new technical approaches and solutions to the field of data mining, but there is still room for further optimization and expansion. Future research will aim to enhance the model's performance, interpretability, and adaptability, promoting the application and development of data mining technologies in a broader range of fields.

Acknowledgments

We thank a bunch of people and funding agency.
Funding: National Key R&D Program of China (Key Special Project for Marine Environmental Security and Sustainable Development of Coral Reefs 2022-3.2)

References

Jing Chen and Yi Zhang. Advances in deep learning for data mining. *Journal of Data Science*, 15 (3):456–478. doi: 10.1007/s10710-017-9314-z.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. doi: 10.1007/s10710-017-9314-z.

Yang Gu, Yu Cheng, Chin-Lung Patricia Chen, and Xia Wang. Proximal policy optimization with policy feedback. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(7):4600–4610, 2022. doi: 10.1109/TSMC.2021.3098451.

Jiawei Han, Michael Kamber, and Jian Pei. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2011. doi: 10.1016/B978-0-12-381479-1.00001-7.

David J Hand. Principles of data mining. *Drug Safety*, 30:621–622, 2007. doi: 10.2165/00002018-200730060-00012.

Ming Li and Hao Zhang. Application of deep learning in data mining. *Journal of Computer Applications*, 41(4):12–18, 2024. doi: 10.1007/s10710-017-9314-z.

Kevin P Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012. doi: 10.1007/s10710-017-9314-z.

Rumen Rafailov, Animesh Sharma, Eric Mitchell, and et al. Direct preference optimization: Your language model is secretly a reward model. In *Advances in Neural Information Processing Systems*, volume 36, pages 53728–53741, 2023. doi: 10.48550/arXiv.2306.08888.

Ritu Rastogi and Monika Bansal. Diabetes prediction model using data mining techniques. *Measurement: Sensors*, 25:100605, 2023. doi: 10.1016/j.measurement.2023.100605.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, and et al. Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008, 2017. doi: 10.48550/arXiv.1706.03762.

Qiang Wang and Lei Zhao. Multimodal data mining techniques in the era of artificial intelligence. *CAAI Transactions on Intelligent Systems*, 19(5):45–56, 2024. doi: 10.1007/s10710-017-9314-z.

Jiachen Yang, Kuan Lin, and Xiaoxiao Yu. Think when you need: Self-adaptive chain-of-thought learning. *arXiv preprint arXiv:2504.03234*, 2025. doi: 10.48550/arXiv.2504.03234.

Qiang Yu, Zi Zhang, Rui Zhu, and et al. Dapo: An open-source llm reinforcement learning system at scale. *arXiv preprint arXiv:2503.14476*, 2025. doi: 10.48550/arXiv.2503.14476.

Wen Zhang, Qiang Wang, Xiang Kong, and et al. Fine-tuning large language models for chemical text mining. *Chemical Science*, 15(27):10600–10611, 2024a. doi: 10.1039/D4SC00000A.

Yi Zhang, Yang Pan, Tian Zhong, and et al. Potential of multimodal large language models for data mining of medical images and free-text reports. *Meta-Radiology*, 2(4):100103, 2024b. doi: 10.1007/s43596-024-00103-9.

Yi Zhang, Ming Zhong, Shuo Ouyang, and et al. Automated mining of structured knowledge from text in the era of large language models. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 6644–6654, 2024c. doi: 10.1145/3646134.3697766.