AE SEMRL: LEARNING SEMANTIC ASSOCIATION RULES WITH AUTOENCODERS

Research by Erkan Karabulut, Victoria Degeler, Paul Groth

Presented by Ronak Urvish Malkan

INTRODUCTION TO ARM AND CHALLENGES

What is ARM?

- Identifies patterns or relationships in data through if-then rules.
- Example: "If a customer buys bread, they are likely to buy butter."

Why is ARM Important?

- Helps uncover hidden patterns in data.
- Used in retail, healthcare, and IoT applications.

Challenges in Traditional ARM:

- Scalability: Struggles with high-dimensional datasets.
- Explainability: Generates too many rules, often lacking context.
- Inefficiency: High execution times for large datasets.

OVERVIEW OF AE SEMRL AND CONTRIBUTIONS

WHAT IS AE SEMRL?

- AE SemRL stands for AutoEncoder-based Semantic association Rule Learning.
- A novel framework that combines autoencoders (a type of neural network) with semantic enrichment to mine association rules.

HOW DOES IT WORK?

- Utilizes semantic knowledge graphs to enrich raw data with contextual information.
- Leverages autoencoders to capture patterns and associations in the latent space.

WHY IS AE SEMRL UNIQUE?

- Combines semantic enrichment (adds context to data) with deep learning (autoencoders) to address traditional ARM challenges.
- Fast and Scalable: Processes high-dimensional data efficiently.
- Interpretable: Generates rules that are meaningful and actionable.

KEY CONTRIBUTIONS

- Semantic Enrichment: Makes rules more generalizable and explainable by embedding domain knowledge.
- Latent Representations: Autoencoders reduce data complexity and extract key patterns.
- Improved Efficiency: Demonstrates hundreds of times faster execution compared to traditional ARM methods.

AUTOENCODER ARCHITECTURE

WHAT IS AN AUTOENCODER?

- A neural network designed to learn compressed representations of input data.
- Comprises two main components:
- Encoder: Compresses the input into a smaller, latent representation.
- Decoder: Reconstructs the input from the latent representation.

DENOISING AUTOENCODER IN AE SEMRL

- Adds noise to the input during training, improving robustness.
- Captures meaningful patterns and associations even in noisy data.

ARCHITECTURE BREAKDOWN

- Input Layer: Accepts semantically enriched data.
- Hidden Layers: Encode features into a compact latent space.
- Latent Layer: Captures the relationships among features.
- Output Layer: Reconstructs the input for validation.

KEY FEATURES

- Robustness: Handles noisy, high-dimensional data.
- Scalability: Effective even with large datasets.
- Interpretability: Encoded features directly relate to association rules.

SEMANTIC ENRICHMENT PROCESS

WHAT IS SEMANTIC ENRICHMENT?

- The process of integrating domain-specific knowledge (metadata) into raw data.
- Provides additional context and meaning, making patterns more interpretable.

HOW DOES SEMANTIC ENRICHMENT WORK IN AE SEMRL?

- Uses knowledge graphs to link raw data with semantic information.
- Enriches numerical or categorical data with contextual properties like:
- Sensor type, location, and environment in IoT datasets.
- Relationships between entities in the data.

WHY IS SEMANTIC ENRICHMENT IMPORTANT?

- Improves Rule Quality: Enriched rules are more generalizable and meaningful.
- Facilitates Explainability: Adds context to associations, making them easier to understand.
- Enables Scalability: Handles high-dimensional data by reducing redundancy through context.

CONCLUSION AND FUTURE DIRECTIONS

SUMMARY OF AE SEMRL'S CONTRIBUTIONS

- Combines semantic enrichment with autoencoders to tackle challenges in Association Rule Mining (ARM).
- Addresses traditional limitations of ARM:
- Scalability: Handles high-dimensional data efficiently.
- Explainability: Generates interpretable, actionable rules.
- Efficiency: Demonstrates hundreds of times faster execution.

FUTURE RESEARCH DIRECTIONS:

Exploring Other Architectures:

- Investigate Graph Neural Networks (GNNs) for learning from graph-structured data.
- Experiment with other deep learning techniques for improved scalability.

Refining Rule Extraction:

- Develop more sophisticated methods to extract rules from latent spaces.
- Incorporate adaptive algorithms to improve rule quality.

Expanding Real-World Applications:

- Apply AE SemRL in real-time systems like anomaly detection and predictive maintenance.
- Extend its use in domains such as financial risk analysis and supply chain optimization.

Thank You

For your attention