



JAGRUTI

Serve Smart Hackathon

FAKE NEWS

FACT
FAKE

PROBLEM STATEMENT

Addressing the Fake News Challenge in the Digital Era

Problem

Fake news spreads rapidly online, influencing public opinion and eroding trust. Traditional methods fail to adapt to evolving language and manipulation techniques.

Key Challenges

- Dynamic, evolving misinformation tactics.
- Limited adaptability of single-model approaches.
- Poor generalization and degrading performance over time.

Solution

Build adaptable, multi-model AI systems for robust detection.

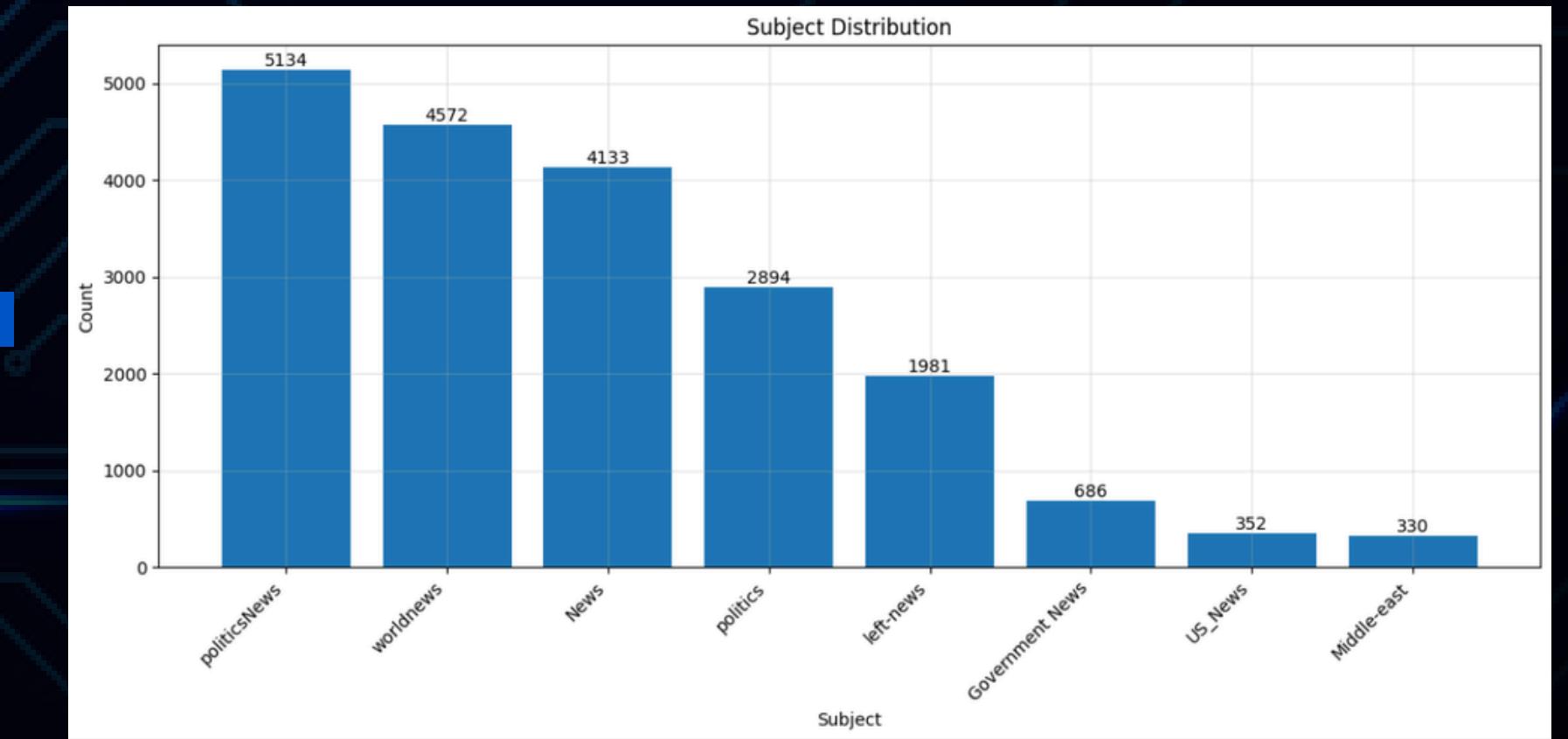
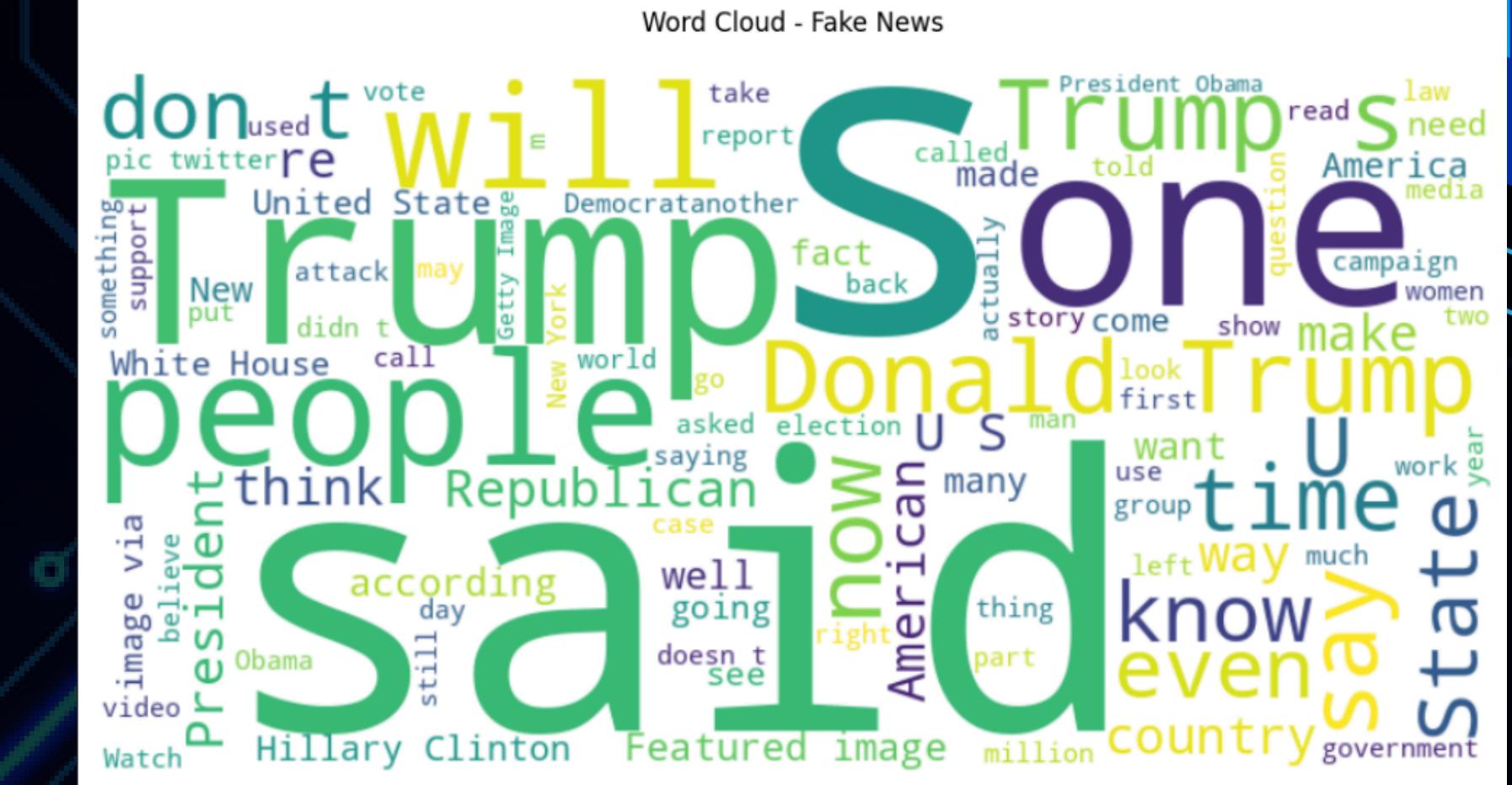
THE DATASET

Columns: Title | Text | Subject | Date | Label

Feature Engineering Approach

- Linguistic Features:
 - Readability metrics for text complexity.
 - Structural patterns: punctuation, sentence length, and document structure.
 - Semantic Features:
 - BERT Embeddings: Captures semantic meaning using advanced tokenization and attention.
 - Dimensionality Reduction: SVD and custom algorithms for optimized representation.

Outcome: Expanded dataset from 5 columns to 7 meaningful features for enhanced detection.



EXPERT 1: CLASSICAL MIL ENSEMBLE MODEL

Model Selection and Optimization:

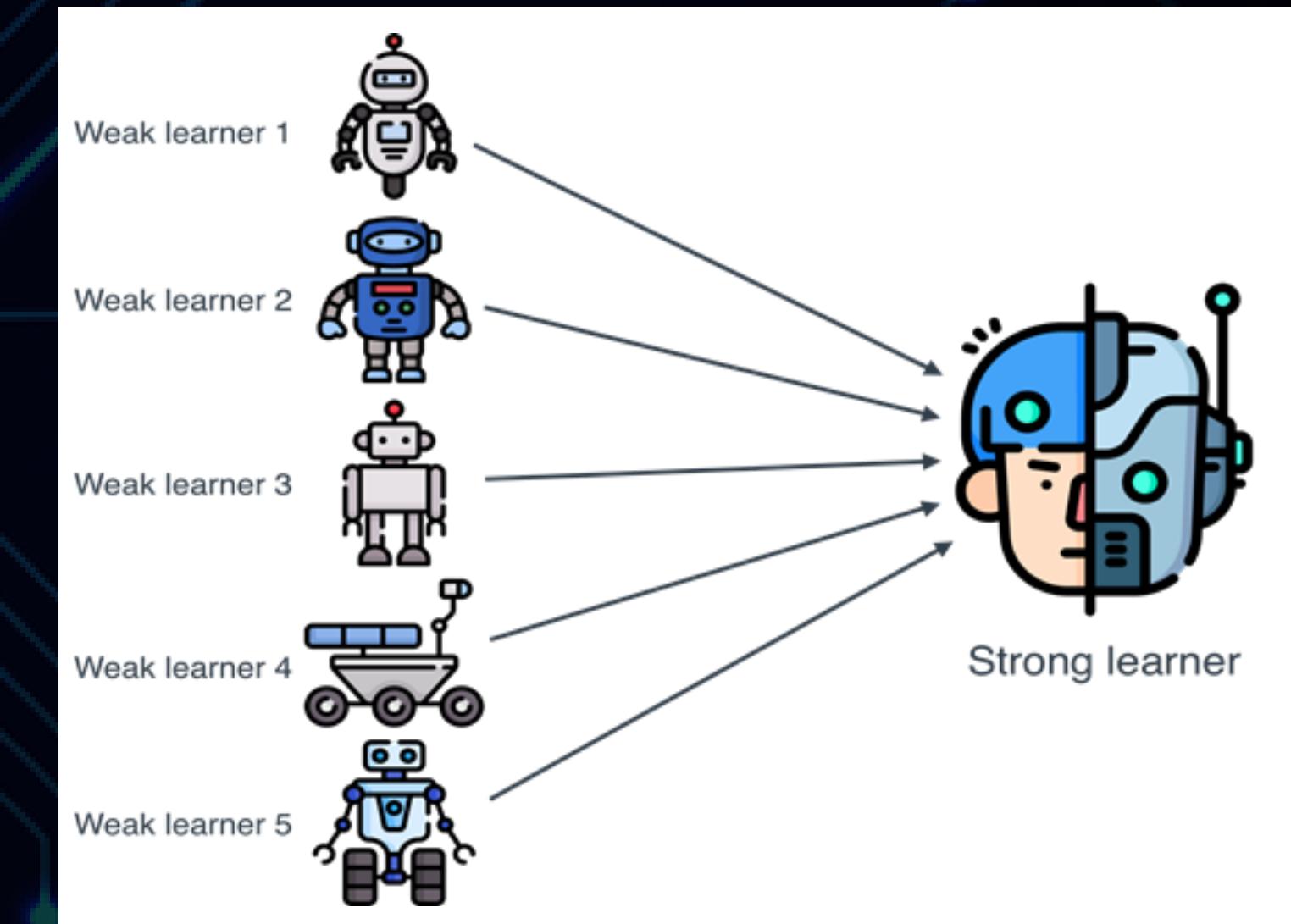
- Gradient Boosting, Random Forest, Logistic Regression, and XGBoost with custom tuning.
- Feature sampling to enhance performance.

Ensemble Integration:

- Weighted voting and cross-validation for robust model integration.

Training Results:

Demonstrated strong reliability and robustness in identifying structural and statistical patterns.



Accuracy	Precision	Recall	F1	Auc Roc
1.0000	1.0000	1.0000	1.0000	1.0000

WHY MOVE BEYOND EXPERT 1?

Perfect Performance in Expert 1:

- Achieved flawless results across evaluation metrics.
- Raised concerns of overfitting: Overly fine-tuned models may lack generalizability.

Adopting MoE:

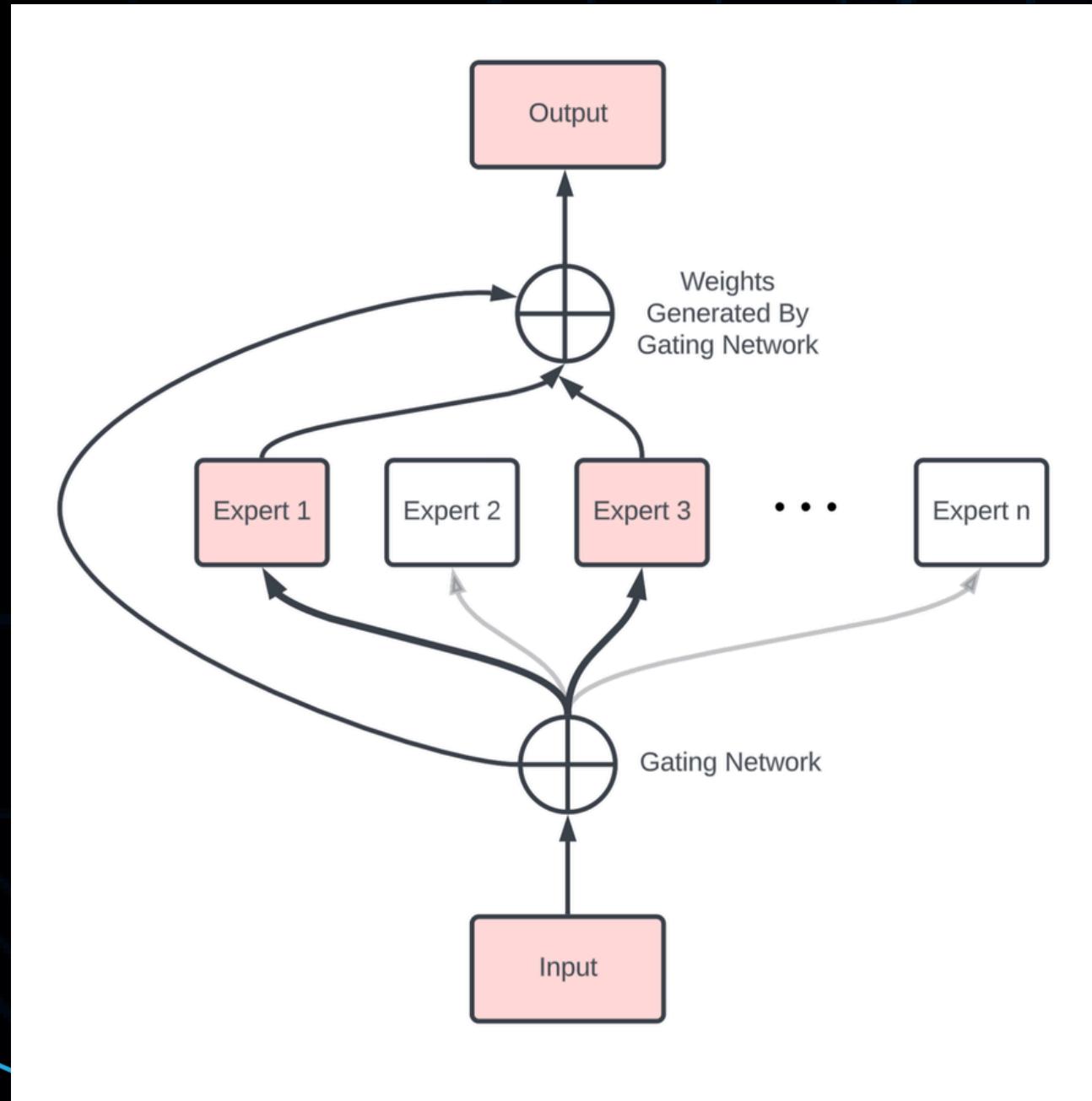
- Incorporating BERT Expert:
 - Fine-tuned BERT transformer added for enhanced adaptability to diverse inputs.
- MoE Advantage:
 - Dynamic Routing: Gating network adapts in real-time, routing data to the most appropriate expert.
 - Flexibility: Unlike static layers, MoE adjusts decision-making based on input characteristics.

Outcome:

- Enhanced Robustness: Prevents overfitting and improves adaptability to emerging misinformation trends.



WHAT IS MIXTURE OF EXPERTS?



Purpose: Enhances model specialization and adaptability by dividing tasks among multiple "experts."

Mechanism:

- Gating Network: Routes input data to the most suitable expert(s) for efficient processing.
- Experts in the System:
 - Classical Ensemble: Identifies structural and statistical patterns.
 - Fine-tuned BERT: Captures semantic nuances and adapts to misinformation trends.

Key Innovation:

Instead of using experts to compensate for each other's flaws, I use them to mitigate overfitting, ensuring robust and accurate predictions.

EXPERT 2: BERT

Custom Fine-Tuning:

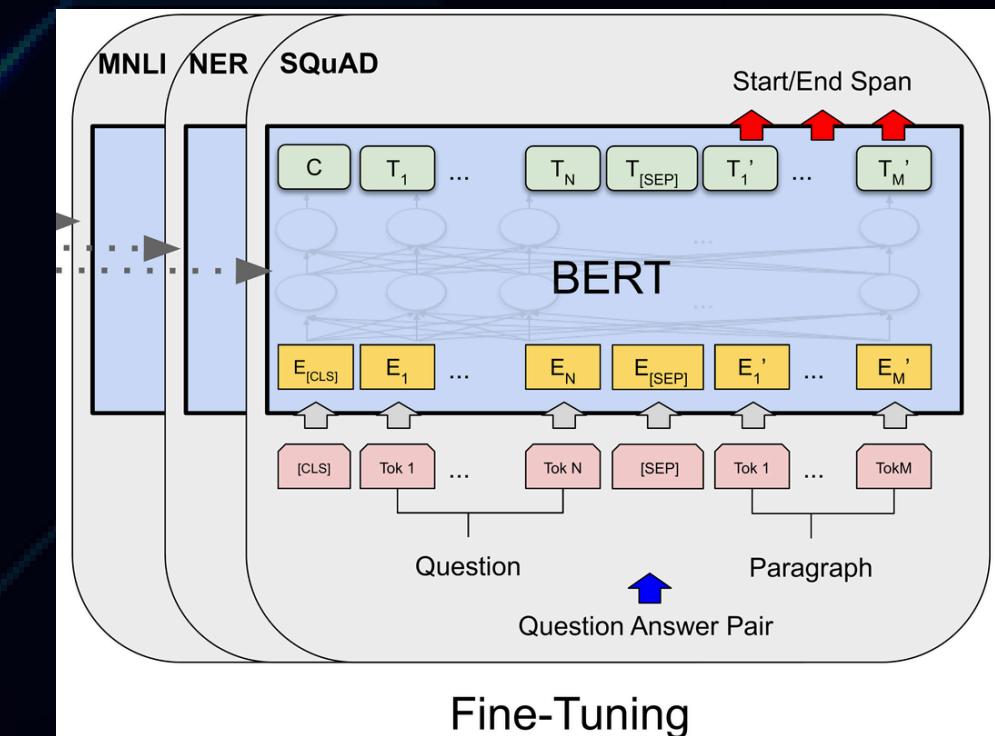
BERT model tailored to the task for better performance on complex textual data.

Training Techniques:

- Gradient Accumulation: Efficiently handles large batch sizes.
- Learning Rate Scheduling: Optimizes training for faster convergence.
- Mixed-Precision Training: Enhances efficiency without compromising model accuracy.

Training Results:

- Performance: Demonstrated strong adaptability to semantic nuances and emerging misinformation trends.



Accuracy	Precision	Recall	F1	Auc Roc
0.9958	0.9982	0.9931	0.9956	0.9996

THE GATE TRAINING

Gating Network Optimization

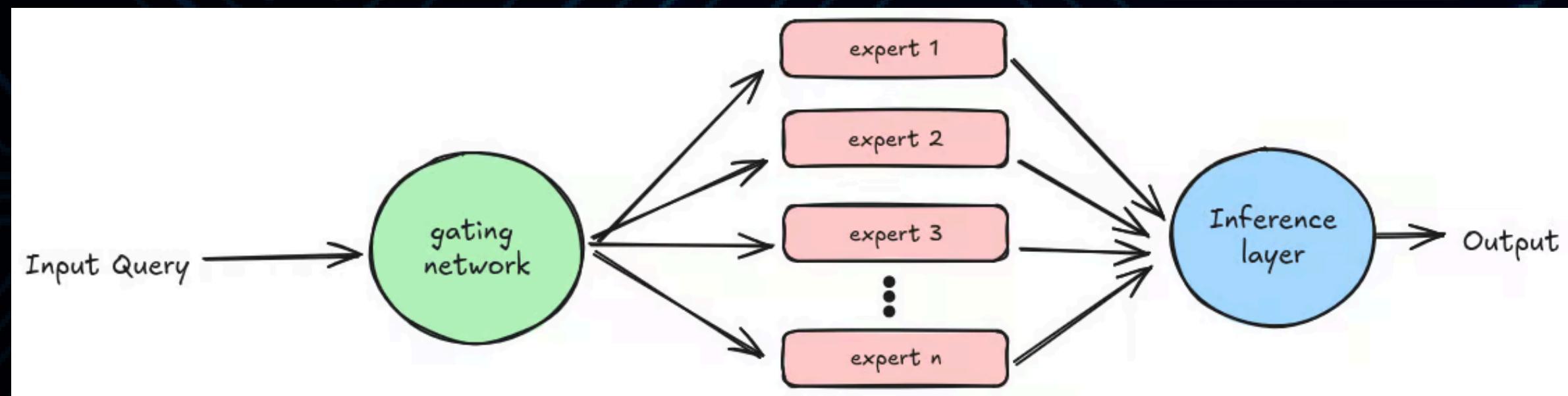
- Temperature-Scaled Softmax & Dropout: Prevents model collapse and ensures diverse expert utilization.
- Loss Penalties: Encourages expert diversity and avoids over-reliance on a single model.

Dimensionality Analysis Insights

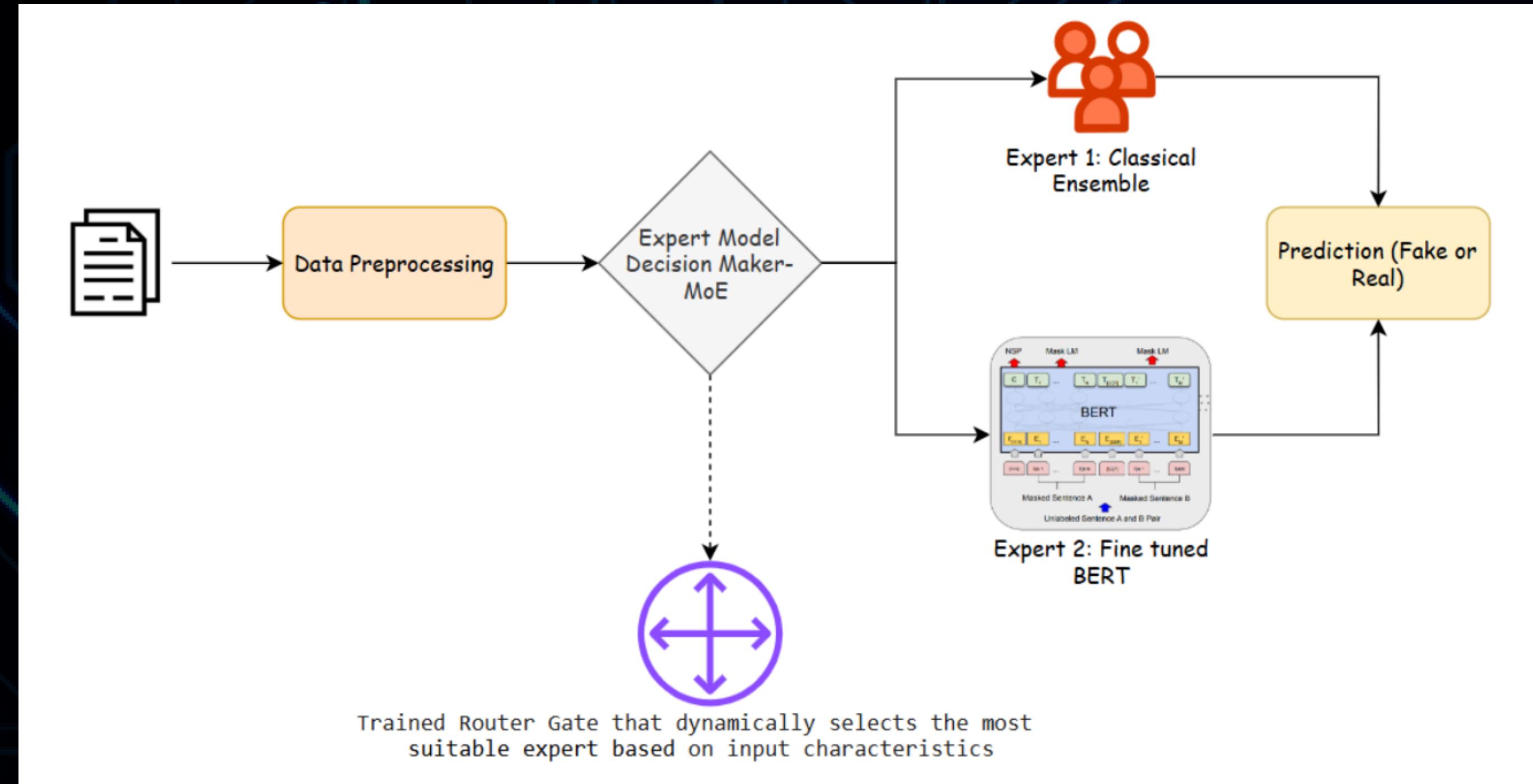
- SVD on BERT Embeddings: Revealed patterns in structural features, semantic relationships, and emotional language.

Temporal Pattern Analysis

- Dynamic Weight Adjustments: Adapted to temporal changes in news content for better performance on emerging sources and breaking events.



ARCHITECTURE



BENCHMARKING

Testing on Other Data

- Dataset: Evaluated on customer dataset generated using the Fake and Real News Dataset from Kaggle.
- Results: Achieved perfect performance metrics, showcasing exceptional accuracy and robustness.

General Performance of Fake News Classifiers

- Typical Challenges:
 - Overfitting, generalization to new trends, and adaptability to evolving misinformation.
- State-of-the-art models often struggle with emerging tactics and require continuous retraining.

Takeaway: The model's performance on both custom and external datasets demonstrates its effectiveness in tackling real-world fake news challenges.

FUTURE SCOPE

Architecture Evolution

- More Experts: Add specialized models for enhanced adaptability.
- Hierarchical Gating & RL: Improve decision-making and system robustness.

Feature Engineering

- Advanced Linguistic Features: Use semantic role labeling for deeper understanding.
- Multimodal Integration: Combine text and other data types for richer analysis.

Performance Optimization

- Distributed & Edge Computing: Improve scalability and efficiency.
- Hardware Acceleration: Use GPUs/TPUs for faster, real-time processing.



THANKS!!

- Sakshi Kumar