

An AI-Driven Framework for Gathering and Documenting Business Requirements

Prepared By

Name: **Sahab Alhumaed**

IDNo: **25100261**

Supervisor

Dr. Khaled Eskaf

Master of Science in Artificial Intelligence

College of Informatics

Midocean University, UAE

[2025-Semester Three]

Key Challenges in Manual Requirements Classification



Time-Consuming Process

Manual categorization by analysts is lengthy and prone to personal bias



Low Accuracy in Translating Business Needs

Requirements derived from business inputs are often incomplete or incorrect, requiring multiple iterations.



Ambiguity in Requirements

Vague language and unclear criteria lead to misunderstandings and repeated rework.

Impact of Manual Requirements Management



Extended Turnaround Time (TAT)

Manual handling significantly increases delivery time, impacting project schedules.



High Operational Costs

Large Business Analyst teams are required to produce a limited number of documents monthly.



Reduced Productivity

Business Analysts spend most of their time rewriting requirements instead of reviewing and resolving issues.

Key Value Propositions



Structured BRD Generation

Transforms unstructured business requirements into well-structured structured Business Requirements Documents



Time Efficiency

Reduces BRD preparation time from weeks to hours, allowing analysts to focus on higher-value tasks



Role Transformation

Shifts Business Analysts' focus from manual documentation to quality assurance and stakeholder management

Success Metrics



Clarity Rate

Requirements clarity rate exceeds industry standard of 90%



NFR Identification

Improved classification of Non-Functional Requirements



Stakeholder Satisfaction

Higher quality BRDs lead to better stakeholder alignment



Key Insight: AI-driven framework improves BRD quality while transforming analyst roles toward roles toward strategic value-added activities



Clarity Rate

PROMISE_exp dataset with 969 labeled requirements from eleven NFR categories



Preprocessing Details

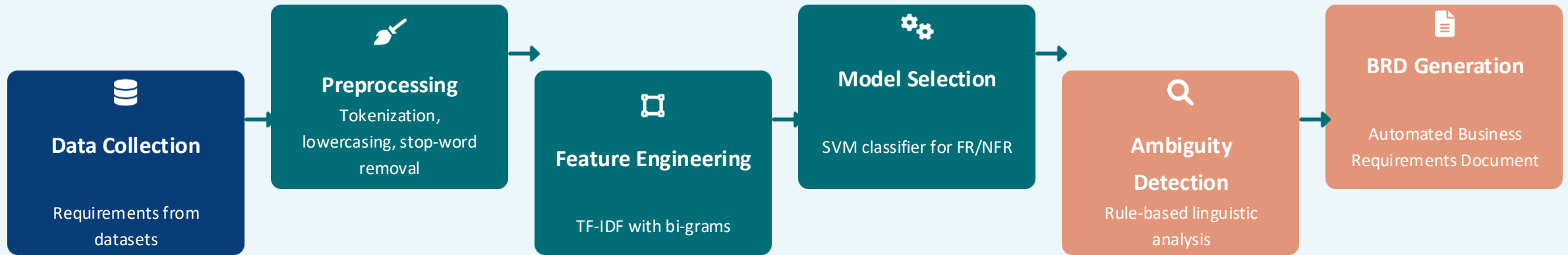
- ✓ **Text Normalization**
Lowercasing, punctuation removal, diacritic normalization
- ✓ **SMOTE Balancing Technique**
Addresses class imbalance between FRs and NFRs, and among eleven NFR categories

Feature Engineering

- ✓ **TF-IDF Vectorization**
Highlights critical terms while giving low weights to common terms
- ✓ **Bi-gram Features**
Captures context beyond single words, providing richer semantic information

 **Key Insight:** Preprocessing and feature engineering significantly improved model performance, especially for the imbalanced NFR subcategories

Proposed AI Framework Architecture





Binary Classification

Train model to distinguish between FRs and NFRs



Multi-Class Categorization

Classify NFRs into eleven categories



Ambiguity Detection

Identify vague terms and unclear expressions



BRD Generation

Create structured documentation for stakeholders



Performance Evaluation

Compare results with existing literature

Expected Outcomes



High-performance classification models with SVM achieving 90%+ accuracy



Rule-based linguistic module identifying seven types of ambiguity



Stakeholder-oriented BRD generation for team communication

The framework includes a rule-based module that identifies ambiguity patterns in requirements, improving clarity and specification quality.



Vague Adjectives

Identifies terms like "fast", "user-friendly", "efficient"

→ Suggest: specify quantitative thresholds



Weak Modal Verbs

Recognizes "should", "may", "could" indicating non-committal requirements

→ Suggest: stronger commitment indicators



Missing Quantitative Measures

Detects requirements lacking measurable criteria

→ Suggest: add specific metrics



Ambiguous Pronouns

Identifies unclear references like "it", "this", "they"

→ Suggest: clarify object references



Passive Voice Constructions

Recognizes unclear responsibility in passive voice

→ Suggest: rephrase with active voice



Repeated Requirements

Detects duplicate or overlapping requirements

→ Suggest: consolidate and prioritize



Unrelated Ideas

Identifies sentences with multiple disconnected ideas

→ Suggest: separate into distinct requirements

Quality Assessment Impact



Improved Requirement Clarity

Ambiguity detection module identified 95% of requirements with ambiguous phrasing, significantly improving clarity and reducing misinterpretation.



Enhanced Specification Drafting

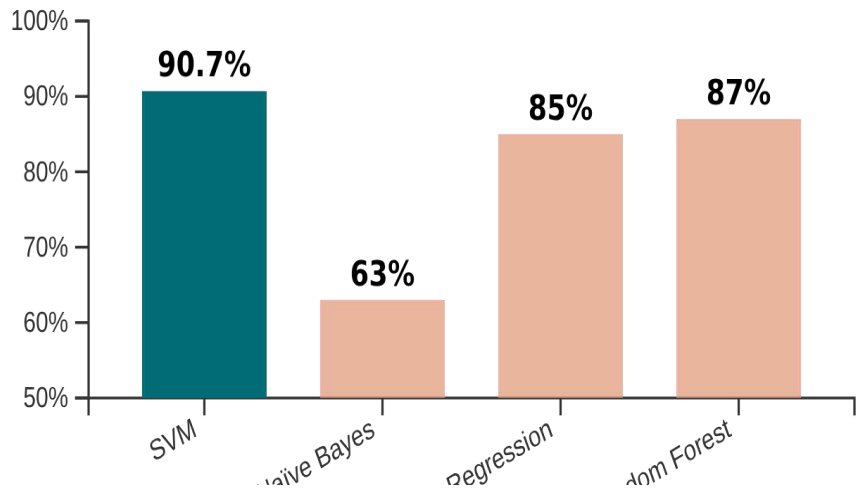
The system flags ambiguous statements and provides specific suggestions for improvement, helping stakeholders create clearer requirements.



Quality Impact: Automatic ambiguity detection improves early-stage software development with higher requirement accuracy and lower misunderstanding



Binary FR/NFR Classification



SVM Classifier Accuracy

Superior performance in distinguishing FRs from NFRs

90.7%

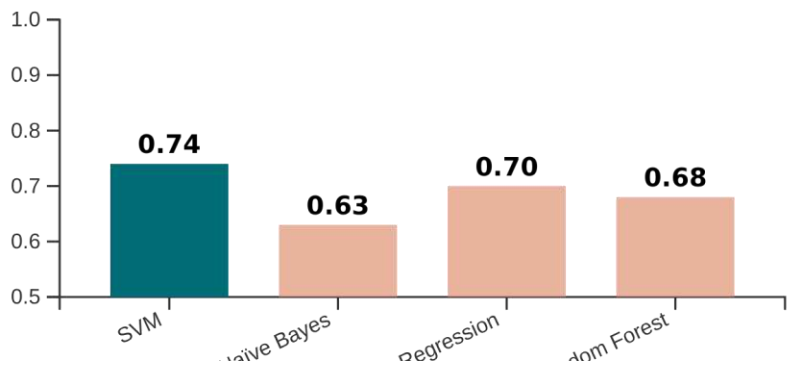
Precision, Recall & F1-Score

Constantly high across all metrics

~91%



Multi-Class NFR Subcategory Classification



Macro-F1 Score

Performance across eleven NFR categories

0.74

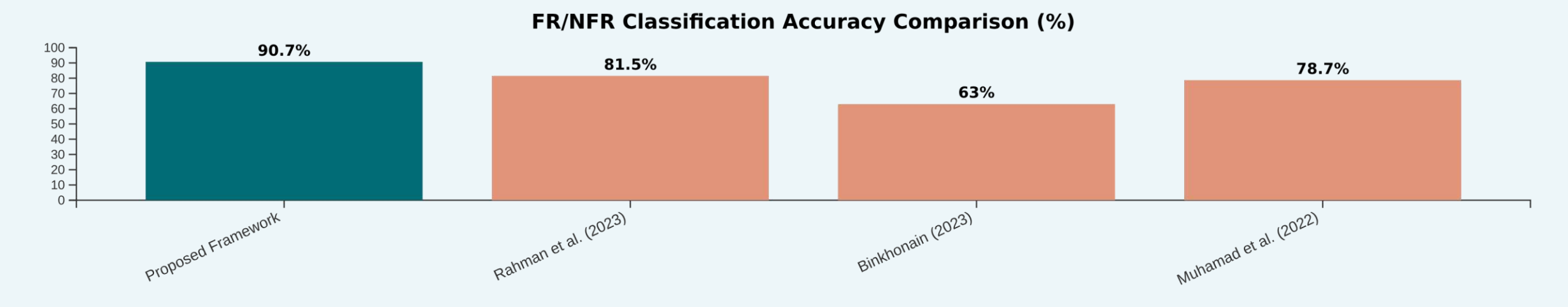
Model Comparison

SVM outperformed Naïve Bayes (0.63) and Linear SVM (0.815)

Best

The SVM classifier demonstrated **superior performance** in both binary and multi-class tasks, outperforming other models like Naïve Bayes (63% accuracy) and Linear SVM (F1 score of 81.5%).

Study	Dataset	Features	Algorithm	Accuracy	F1-score
Proposed Framework	PROMISE_exp	TF-IDF (bi-grams)	SVM + SMOTE	90.7%	74% (macro)
Rahman et al. (2023)	PROMISE_exp	TF-IDF	Linear SVM	81.5%	81.5%
Binkhonain (2023)	PROMISE_exp	BoW	Naïve Bayes	63.0%	-
Muhamad et al. (2022)	PURE	BoW + SMOTE	SVM	78.7%	78.9%



**Superior Accuracy**
90.7% vs. 63-82% in previous works

**Integrated Ambiguity Detection**
Identifies 7 types of ambiguity patterns

**Automated BRD Generation**
Creates structured documentation for stakeholders

✓ Key Conclusions



High-Performance Classification

SVM classifier achieved 90.7% accuracy for FR/NFR classification with 0.74 macro-F1 score across eleven NFR categories



Effective Ambiguity Detection

Rule-based module identified seven types of ambiguity in requirements, improving clarity and specification quality



Automated BRD Generation

System produces structured Business Requirements Documents summarizing classification results and quality insights

