（Form6）

論　文　要　旨
Thesis Abstract

(yyyy/mm/dd)　2025　年　06　月 18 日

※報告番号	甲第３７７ 号	氏　名 (Name)	GLORIANA JOSEPH MONKO

主論文題名　（Title）

A Multi-Stage Framework for Adverse Drug Reaction Detection: Integrating SS-DBSCAN Clustering, Human-in-the-Loop for Labeling, and Transformers

内容の要旨　（Abstract）

Adverse Drug Reactions (ADRs) pose significant threats to patient safety and public health, yet current pharmacovigilance systems continue to face critical limitations, including under-reporting, poor data quality, and a suitable framework for detecting ADRs. With the increasing availability of Electronic Medical Records (EMRs), there is a growing opportunity to automate ADR detection using advanced machine learning and natural language processing techniques. This thesis presents a robust and scalable framework for ADR detection from high-dimensional clinical text data, integrating unsupervised clustering, semi-supervised learning, and supervised learning with transformer-based models.

To address the complexity and noise inherent in EMR data, an enhanced clustering algorithm, Stratified Sampling DBSCAN (SS-DBSCAN), was developed. This method employs a stratified sampling strategy for dynamic epsilon estimation and a fast grid search technique to optimize the MinPts parameter, thereby improving cluster quality and interpretability.

A human-in-the-loop (HITL) mechanism was incorporated to refine cluster labels from SS-DBSCAN by iteratively fine-tuning the transformer BERT model, thereby enhancing labeling accuracy through expert label feedback. This approach introduces a novel

technique for data labeling. Using this approach, we efficiently labeled 50,000 samples from the MIMIC III dataset, demonstrating the scalability and reliability of our framework in producing high-quality training data for supervised learning.

A total of 60,000 labeled datasets from MIMIC III and ADE_classification data were then used to fine-tune three transformer-based models (BERT, GPT-2, and LLaMA 3 (1B)) for ADR classification. The models' performances were evaluated using both validation and confusion metrics. BERT achieved the highest validation accuracy of approximately 99%, demonstrating excellent generalization. GPT-2 and LLaMA 3 reached a validation accuracy of 98.8%, GPT-2 with fewer false negatives than all other models, and BERT with fewer false positives. All three models demonstrated near-perfect training accuracy, confirming their effectiveness in extracting meaningful patterns from semi-supervised clinical data.

The proposed framework demonstrates robust performance in high-dimensional, noisy datasets, significantly advancing data labeling and ADR detection. By combining SS-DBSCAN, HITL, and transformer-based models, this research contributes a reliable, interpretable, and practical solution to modern pharmacovigilance challenges.