

A Teacher-Student BERT Architecture for Semi-Supervised Learning on Unstructured Social Work Texts: Identifying Service Gaps and Needs

Yih-Chang Chen*

Department of Social Work, Chang Jung Christian University, Tainan 711, Taiwan, China

E-mail: cheny@mail.cjcu.edu.tw

ORCID iD: <https://orcid.org/0000-0002-6149-6228>

*Corresponding author

Chia-Ching Lin

Department of Finance, Chang Jung Christian University, Tainan 711, Taiwan, China

E-mail: lincc@mail.cjcu.edu.tw

ORCID iD: <https://orcid.org/0000-0003-1689-6131>

Sedat Agan

Department of Physics, Kirikkale University, Kirikkale 71450, Turkey

E-mail: sedatagan@hotmail.com

ORCID iD: <https://orcid.org/0009-0002-3497-3444>

Received: April 10, 2026; Revised: June 7, 2026; Accepted: July 14, 2026; Published: October 8, 2026

Abstract: To address the “information gap” in social work arising from unstructured data and limited labeled instances, this study proposes a semi-supervised learning framework based on a Teacher-Student BERT architecture. Based on an analysis of 50,000 case records spanning 2019 to 2024, the model incorporates Latent Dirichlet Allocation for topic extraction alongside a multidimensional gap analysis. The proposed method attained a state-of-the-art F1 score of 0.913, markedly surpassing baseline BERT models. Notable findings include a 67.6% increase in mental health-related discourse and the quantification of significant systemic deficiencies, such as inadequate coverage in elderly care (45.8%) and substantial unmet needs in medical assistance (46.4%). Furthermore, to address the profound challenges of class imbalance and the data-hungry nature of transformer models, this study integrated generative artificial intelligence (AI) data augmentation techniques. This approach produced synthetically varied case narratives that preserved the original socio-economic context while expanding lexical and structural diversity, thereby significantly enhancing the classification accuracy of minority classes. Additionally, the use of a Mean Teacher denoising framework and knowledge distillation drastically reduced the computational inference time, rendering the architecture highly suitable for deployment in resource-constrained social welfare environments. Prior to analysis, all case records underwent rigorous automated Personally Identifiable Information (PII) scrubbing utilizing the Microsoft Presidio framework to guarantee data privacy. To ensure reproducibility, the Teacher-Student training codebase, prompt architectures, and anonymized synthetic data samples will be made publicly available upon request. This research effectively transforms administrative textual data into actionable strategic intelligence, offering a scalable and evidence-based tool to enhance resource allocation and inform policy development.

Index Terms: Semi-Supervised Learning, Natural Language Processing (NLP), Social Work, Service Gap Analysis, Topic Modeling, Knowledge Distillation, Data Augmentation.

1. Introduction

1.1. Problem Statement and Research Context

The discipline of social work has persistently grappled with a critical issue: the effective extraction of salient information from extensive case records to enhance decision-making and optimize service delivery. This challenge

fundamentally stems from the “information gap” — a severe discrepancy between the vast amounts of localized, unstructured administrative records generated daily and the structured, macro-level insights required for immediate policy formulation and targeted intervention. Traditionally, social workers have depended on manual case review and experiential judgment to assess client needs. This approach is not only labor-intensive and time-consuming but also lacks scalability for systematic analysis. The rapid expansion of social welfare systems and the increasing heterogeneity of service recipients have led to an exponential growth in case records, rendering conventional manual analysis methods insufficient to meet the demands of contemporary social work practice [1-2].

Case records, as fundamental documents within social work practice, encapsulate essential information about service recipients, problem descriptions, intervention processes, and service outcomes. These textual data harbor rich insights, including the nature of social problems encountered by clients, evolving service needs, and the extent of coverage within existing service systems. However, the predominantly unstructured textual format of case records, coupled with individualized writing styles and the complexity of specialized terminology, poses significant challenges for traditional data analysis techniques [1,3]. The lack of standardization in how social workers document psychological distress, economic instability, and domestic strife creates a high-dimensional, sparse data environment that confounds standard analytical procedures.

Recent advancements in artificial intelligence, particularly in natural language processing (NLP) and machine learning, have introduced powerful methodologies for processing unstructured textual data [4-6]. With AI increasingly permeating various aspects of society, including healthcare and welfare management, the adoption of the Transformer neural network architecture is rapidly revolutionizing the processing of unstructured clinical and administrative records [7]. Nonetheless, the application of these technologies in social work remains fraught with challenges: firstly, the acquisition of professionally annotated datasets is resource-intensive, necessitating expert social workers to undertake extensive manual labeling; secondly, the inherent complexity and heterogeneity of social work cases limit the efficacy of purely supervised learning approaches; and finally, extant research predominantly concentrates on identifying isolated issues, lacking comprehensive analyses of systemic service gaps [8-10].

Furthermore, generating a sufficiently large amount of training data for robust information extraction in this domain is either financially prohibitive or runs the risk of introducing highly noisy instances if distant supervision is used indiscriminately [11]. Empirical evidence regarding the deployment of these advanced models in authentic, dynamic, and highly unpredictable human-service environments remains alarmingly scarce, representing a critical limitation in understanding the practical applicability of Transformer-based systems [12].

1.2. Research Objectives and Significance

This study endeavors to develop a semi-supervised learning-based analytical system for social work case records that facilitates automated topic extraction and service gap analysis, thereby furnishing empirical support for policy formulation and service management. The principal research objectives are structurally divided to ensure both technological innovation and practical socio-economic applicability.

Primary Objectives:

- To design a semi-supervised learning model tailored to social work case records for the automatic classification and identification of multiple social issues.
- To establish a topic modeling framework aimed at uncovering emergent and latent social problems embedded deep within unstructured narratives.
- To devise a quantitative service gap analysis methodology to evaluate the coverage status of existing social welfare systems.
- To construct a comprehensive decision support system that delivers real-time, data-driven insights for policymakers and frontline service administrators.

Secondary Objectives:

- To compare the efficacy of various machine learning algorithms in the context of social work text analysis.
- To validate the advantages of semi-supervised learning approaches in environments characterized by limited labeled data.
- To assess the extent to which generative AI-based data augmentation techniques can mitigate the negative impacts of class imbalance within specialized social work datasets.
- To investigate the applicability of topic modeling techniques in detecting cross-domain social issues and overlapping psychosocial stressors.
- To analyze the spatial and temporal dynamics of identified service gaps.
- To evaluate the internal mechanics of the Teacher-Student architecture, explicitly analyzing how linguistic skills and semantic understanding are compressed and transferred from a massive pre-trained teacher to an operationally efficient bidirectional student model.

Theoretically, this research contributes by introducing semi-supervised learning methodologies into the social work domain, thereby broadening the application of artificial intelligence within social science research. It also proposes an

integrated analytical framework that synergistically combines topic mining with service gap analysis, offering novel methodological tools for social policy investigation. Furthermore, through the in-depth mining of large-scale textual data, the study aims to reveal patterns and trends in social problems that may elude traditional research methods.

Practically, the findings of this study hold direct applicability in social work practice: enhancing social workers' capacity to identify and classify case issues efficiently; providing institutional managers with decision support for forecasting service demands and allocating resources; assisting policymakers in pinpointing deficiencies within service systems to inform targeted policy interventions; and fostering inter-agency information sharing and collaboration to optimize the overall efficacy of social welfare operations.

1.3. Literature Review

A. Applications of Natural Language Processing in Social Work

Recent years have witnessed the maturation of NLP technologies within the social work sector. Dorr et al. pioneered investigations into the feasibility of employing NLP techniques to detect social needs, such as chronic stress, social isolation, economic insecurity, and housing instability, from clinical notes authored by primary care physicians [13]. Conway et al. developed the Moonstone system, which utilizes rule-based classification to identify housing status and social support indicators within electronic health records [14].

Concerning the classification of social work interventions, prior studies have constructed classification schemas encompassing ten categories of social work interventions and applied NLP alongside machine learning algorithms to categorize interventions documented in electronic health records. Bako et al. demonstrated that high-performing classifiers, including core support vector machines (accuracy 0.97), logistic regression (accuracy 0.96), and linear support vector machines (accuracy 0.95), exhibit excellent performance in this domain [1]. However, these rule-based and standard machine learning approaches often fail to capture the deep bidirectional context inherent in complex case narratives.

B. Advancements in Semi-Supervised Learning within Healthcare

Semi-supervised learning has experienced rapid progress in healthcare applications. In mental health monitoring, Casalino et al. illustrated that semi-supervised algorithms surpass supervised counterparts in predicting bipolar disorder episodes, particularly under conditions of limited labeled data [8].

In medical image segmentation, semi-supervised frameworks grounded in foundational models, such as SemiSAM+, have achieved substantial performance gains despite minimal annotation availability [15]. In the realm of textual information extraction, semi-supervised learning has emerged as a cost-efficient solution to combat the lack of training data. Recent adaptations of the Mean Teacher framework, a denoising semi-supervised learning methodology, have demonstrated that it is possible to extract highly complex semantic relations between entities while nearing the performance of fully supervised approaches using only 10% of a labeled corpus [16]. This framework's ability to minimize confirmation bias through consistency regularization provides a critical methodological underpinning for the present study's approach to social work records.

C. The Evolution of Transformer Architectures and Knowledge Distillation

While Transformer models like BERT have become the gold standard for natural language understanding, they are notoriously computationally expensive and slow for practical, high-throughput applications. In public welfare and social work institutions, which often operate on constrained IT budgets, deploying massive language models is practically unfeasible. This limitation has inspired significant recent efforts in model compression via knowledge distillation. In this paradigm, a large, robust BERT model acts as a "Teacher," transferring its generalized latent representations and predictive probability distributions to a much smaller "Student" model [17-18]. Extant literature indicates that such architecturally diverse students, ranging from down-scaled BERT variants to bidirectional LSTM models, can be up to 14,000 times smaller and 1,100 times faster than their teachers [19]. Despite these impressive compression rates, research assessing how effectively these student models retain nuanced semantic understanding when faced with difficult, domain-specific linguistic tasks remains sparse. This study bridges this gap by rigorously evaluating the linguistic capabilities retained by the student model within the highly specialized context of social work discourse.

D. Integration of Large Language Models (LLMs) and Generative Data Augmentation

The advent of advanced Large Language Models (LLMs) has introduced unprecedented capabilities in handling unstructured text. Recent research underscores the utility of LLMs, such as Mistral 7B and advanced generative BERT variations, not just for classification, but for generating dynamic embeddings derived from historical interaction patterns [20-22]. In educational and cognitive domains, models like MCQStudentBert have leveraged LLMs to map the contextual history of individuals, providing granular support and forecasting future behavioral patterns [23]. Concurrently, generative AI has been highly effective in data augmentation. To mitigate data scarcity and enhance model robustness against severe class imbalances, researchers have successfully employed LLMs to produce synthetically varied text that retains core pedagogical or social intents while expanding lexical diversity [24]. The integration of these generative augmentation strategies into the semi-supervised pipeline represents a critical advancement over traditional static NLP

approaches.

E. Utilization of Topic Modeling in Social Work Note Analysis

Topic modeling methodologies have yielded promising outcomes in the analysis of social work notes. A large-scale investigation analyzing 950,000 social work notes employed Latent Dirichlet Allocation (LDA) to extract eleven topics pertinent to health-related social determinants, including economic status, abuse history, social support, mortality risk, and mental health [3].

Sun et al. confirmed that topic modeling effectively uncovers latent themes, identifies embedded key information within notes, and facilitates systematic comparisons of topic distributions across diverse note types [3]. These findings substantiate the adoption of topic modeling-based approaches for topic extraction in the present study.

F. Theoretical and Practical Foundations of Service Gap Analysis

Service gap analysis constitutes a vital instrument for evaluating social policies and has been implemented across various countries and regions [25-26]. Joosse et al. conceptualized service gaps at three distinct levels: awareness gaps (where residents lack knowledge of social service providers or referral programs), acceptance gaps (where residents do not accept effective social service referrals), and availability gaps (where residents are unable to identify or promptly contact agencies offering necessary services) [27].

This multidimensional framework offers a robust theoretical foundation for developing quantitative evaluation methods within this study. When integrated with text analysis techniques, it enables more precise identification and quantification of diverse service gap types, transforming abstract theories into measurable empirical realities.

The remainder of this paper is organized as follows: Section 2 details the dataset characteristics, privacy-preserving methodologies, and the technical formulation of the proposed Teacher-Student architecture. Section 3 presents the quantitative evaluation, service gap analysis, and a discussion of ethical implications and limitations. Finally, Section 4 concludes the paper and outlines prospective directions for future research.

2. Materials and Methods

2.1. Research Design and Methodological Framework

This investigation utilizes a mixed-methods strategy, integrating quantitative text analytics with qualitative content analysis to develop a comprehensive framework for examining social work case records. The research design encompasses six principal phases: data acquisition, preprocessing, expert annotation, model training, evaluation and validation, and the application of service gap analysis. The methodological framework is meticulously structured to ensure reproducibility and rigor, as summarized in Table 1.

Table 1. Research methodological framework

Stage	Description	Technical Method	Estimated Time (weeks)
Data Collection	Collect social work case record text data	Database query, API interface	2
Preprocessing	Text cleaning, word segmentation, stop word removal, generative data augmentation	NLTK, Jieba word segmentation tool, LLM Synthesis	3
Annotation	Professional social workers annotate core issue categories	Expert annotation, active learning	8
Model Training	Semi-supervised learning model training	BERT, BiLSTM, SSL, Mean Teacher Distillation	12
Evaluation	Model performance evaluation and validation	Cross-validation, Confusion matrix	4
Application	Service gap analysis and policy recommendations	Statistical analysis, Visualization	6

The entire research timeline is projected to span 35 weeks, with the model training stage identified as the most time-intensive, reflecting the inherent complexity of deep learning architectures and their substantial computational demands. Although the expert annotation phase requires eight weeks, it is indispensable for generating a high-quality labeled dataset required to initialize the Teacher model.

An overview of the methodological workflow is depicted in Fig. 1.

2.2. Data Collection and Sources

Data Sources

The dataset employed in this study is primarily derived from the case management systems of three prominent social welfare organizations, encompassing social work case records from 2019 through 2024. To ensure representativeness and heterogeneity, the selected institutions include: a comprehensive social service center, a shelter specializing in domestic violence cases, and a rehabilitation center dedicated to drug dependency treatment.

Data Scale and Characteristics

The compiled dataset comprises approximately 50,000 case records, amounting to roughly 8 million Chinese characters. Each record averages 160 characters and contains sections such as basic client information, problem assessment, service plans, intervention process documentation, and case closure summaries. The data span ten principal categories of social issues, including substance abuse, domestic violence, financial hardship, child protection, elderly care, mental health, employment assistance, housing difficulties, educational support, and medical aid.

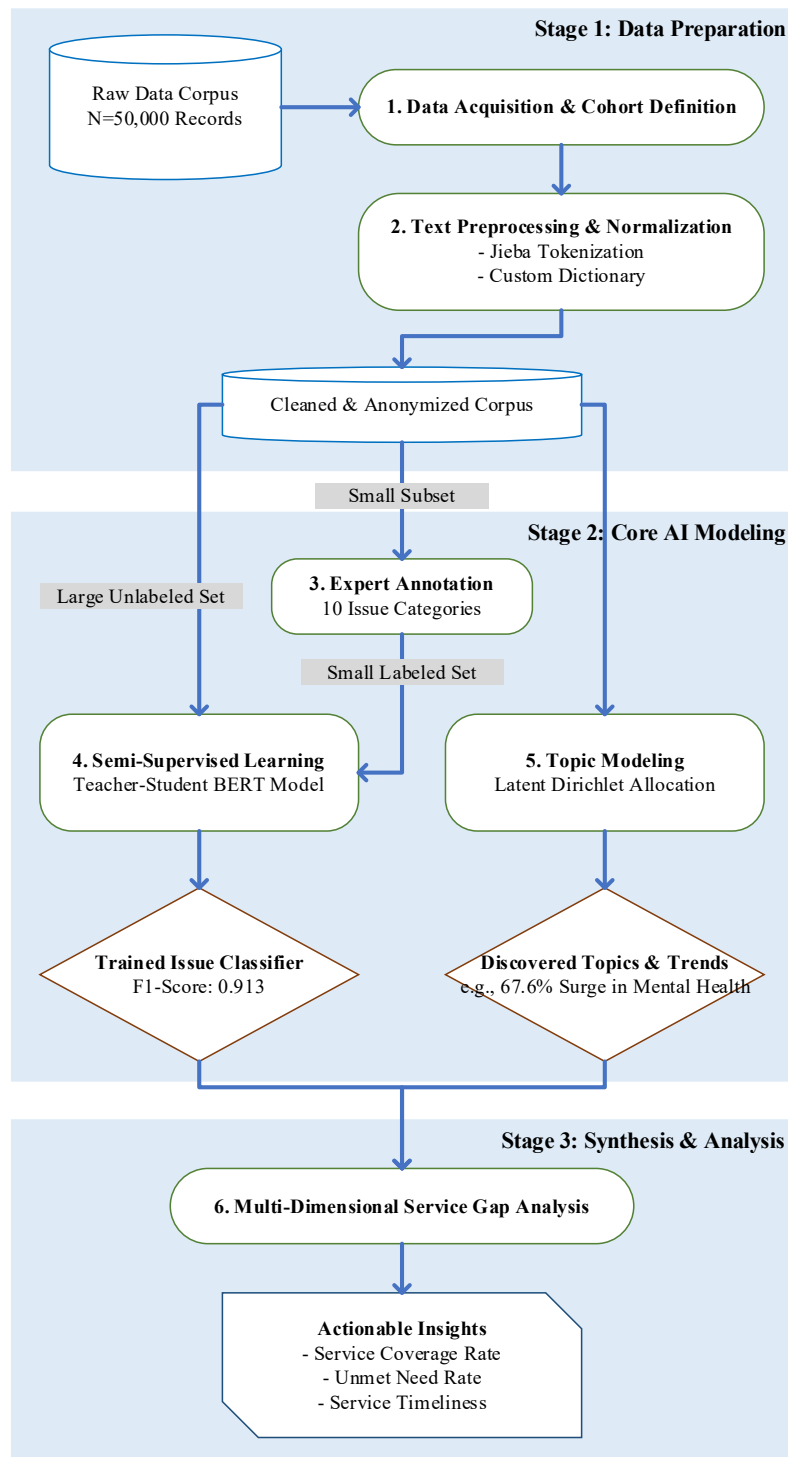


Fig. 1. The methodological pipeline

Within this extensive corpus, exactly 5% (2,500 records) were rigorously annotated by expert social workers to serve as the high-quality supervised training subset, while the remaining 95% (47,500 records) constituted the unlabeled dataset utilized for the semi-supervised consistency training. To accurately account for the temporal concept drift induced by the macro-level societal shifts during the COVID-19 pandemic, the train/test splits were chronologically stratified rather than

purely randomized, ensuring robust temporal representation.

Data Quality Control

To maintain data integrity, stringent quality control procedures were implemented: records shorter than 50 characters were excluded due to insufficient informational content; entries exhibiting excessive duplication or formatting anomalies were removed; and records lacking essential fields (e.g., basic client details or problem descriptions) were omitted.

Most critically, given the highly sensitive nature of social work records, a comprehensive privacy-preserving strategy was enacted. All 50,000 records were processed through an automated anonymization pipeline utilizing the Microsoft Presidio framework. This step successfully stripped all Personally Identifiable Information (PII), including proper names, localized geographic markers, telephone numbers, and specific institutional identifiers, thereby guaranteeing complete de-identification prior to any model exposure.

2.3. Text Preprocessing and Feature Engineering

Text Cleaning and Normalization

Text preprocessing constituted a critical step to optimize model performance. Initially, raw textual data were cleansed by eliminating HTML tags, special characters, and extraneous whitespace. Given the specific nature of social work case records, customized normalization protocols were devised to standardize prevalent abbreviations, domain-specific terminology, and formatting inconsistencies.

Chinese Text Segmentation and Part-of-Speech

Chinese text segmentation was performed using the Jieba tool, augmented with a specialized dictionary tailored to the social work domain to ensure precise segmentation of technical terms. A custom lexicon comprising 2,000 professional social work terms was developed, encompassing various social issues, intervention strategies, and assessment instruments.

Stop Word Processing and Feature Selection

A stop word list specifically designed for social work texts was constructed, incorporating both general linguistic stop words and domain-specific high-frequency non-informative terms. Feature selection was conducted using the term frequency-inverse document frequency (TF-IDF) method, preserving salient features relevant to classification tasks. Ultimately, a feature dictionary containing 15,000 terms was established.

Handling Class Imbalance via Generative AI Data Augmentation

A preliminary analysis of the initial dataset revealed a severe class imbalance. Categories such as “Domestic Violence” and “Financial Hardship” were heavily represented, whereas critical but less frequently documented issues like “Educational Support” and “Housing Difficulties” constituted less than 5% of the total corpus. Given the data-hungry nature of Transformer models, this imbalance posed a significant threat to classification accuracy, potentially causing the model to overfit majority classes while ignoring minority ones [12].

To mitigate this potential data scarcity and enhance overall model robustness, advanced generative AI techniques were employed for data augmentation. Synthetic case records were generated by utilizing a pre-trained LLM architecture instructed with highly specific, domain-bounded prompts. The generative model was constrained to produce synthetically varied utterances that strictly retained the original social work intent, professional tone, and diagnostic structures while expanding the lexical and syntactic diversity of the minority classes. This process iteratively expanded the minority datasets until a balanced class distribution was achieved, ensuring that the subsequent Teacher-Student training regimen received adequate exposure to all ten principal categories.

2.4. Semi-Supervised Learning Model Design

Model Architecture Design

The fundamental machine learning task is formulated as a multi-class classification problem designed to categorize unstructured client narratives into one of the ten principal social work issue categories. The proposed semi-supervised learning framework employs a Teacher-Student architecture, integrating the BERT pre-trained language model with topic modeling methodologies [28-29]. The model comprises three principal components: a feature extraction module, a semi-supervised learning module, and a topic modeling module.

To combat the introduction of noisy instances during distant or weak supervision, this study adapted the Mean Teacher paradigm, a highly effective denoising semi-supervised framework. In this setup, the “Teacher” model maintains an exponential moving average of the “Student” model’s weights. The Student model learns from both the heavily augmented supervised dataset and the pseudo-labels generated by the Teacher model on the vast array of unlabeled case records. Furthermore, syntax-aware representations were incorporated into the neural framework to enhance the learned semantic embeddings of complex, multi-clause Chinese sentences.

Mathematical Model Representation

The semi-supervised learning objective function is formulated as follows:

$$L = L_s + \lambda \cdot L_u + \gamma \cdot R \quad (1)$$

where:

- L_s denotes the supervised cross-entropy loss computed from the expert-annotated and synthetically augmented labeled data;
- L_u represents the unsupervised consistency loss derived from the unlabeled data. Specifically, L_u utilizes the Mean Squared Error (MSE) between the predictive probability distributions of the Student model and the Mean Teacher model, penalizing any divergence when exposed to perturbed versions of the same unlabeled input;
- R is a regularization term designed to mitigate overfitting;
- λ, γ are dynamically adjusted hyperparameters weighting the contribution of each loss component throughout the training epochs.

BERT Feature Extraction

The BERT encoder is mathematically expressed as:

$$H = \text{BERT}(E) \quad (2)$$

where E corresponds to the input word embedding matrix, and H signifies the resulting contextualized representation matrix.

To further refine these representations, historical interaction embeddings, which capture the longitudinal sequence of a client's past engagements with the welfare system, were generated via a bidirectional LSTM layer and concatenated with the contextualized BERT output, creating a multidimensional temporal-semantic feature vector [30].

Classifier Design and Knowledge Distillation

The final classifier produces a probability distribution defined by:

$$P(y | x) = \text{soft max}(W_c \cdot h + b_c) \quad (3)$$

where h is the hidden representation output by BERT, and W_c and b_c denote the weight matrix and bias vector of the classification layer, respectively.

The distillation process forces the computationally efficient Student model to replicate the intricate decision boundaries formed by the massive Teacher network, effectively preserving the semantic acuity of large-scale Transformers while drastically reducing the operational latency.

2.5. Topic Modeling and Issue Mining

Implementation of the LDA Topic Model

The Latent Dirichlet Allocation (LDA) model is used for topic modeling, mathematically represented as:

$$P(W, z, \Theta, \Phi | \alpha, \beta) = \prod_{d=1}^M P(\theta_d | \alpha) \prod_{n=1}^{N_d} P(z_{d,n} | \theta_d) \cdot P(w_{d,n} | \phi_{z_{d,n}}, \beta) \quad (4)$$

where w represents observed words, z represents latent topic assignments, θ denotes the document-topic distribution governed by the Dirichlet prior α , and ϕ denotes the topic-word distribution governed by the Dirichlet prior β . Model parameters are learned through a variational inference algorithm to uncover latent topic structures.

While modern embedding-based topic modeling techniques such as BERTopic or Top2Vec seamlessly integrate with dense Transformer architectures, LDA was explicitly retained and theoretically justified for this specific study. In clinical social work, practitioners rely heavily on exact, granular lexical choices (e.g., specific diagnostic jargon, risk markers, or statutory terms) to formulate safety assessments. LDA's classical bag-of-words generative approach preserves this direct, transparent term-frequency mapping, offering a high degree of interpretability that dense contextual embeddings occasionally obscure [3]. Furthermore, LDA provides a computationally lightweight diagnostic baseline that remains highly familiar to traditional public health stakeholders.

Determination of Optimal Topic Number

An integrative approach combining coherence scores (c_v) and perplexity measures is utilized to identify the optimal

number of topics. A grid search is performed over a range of 5 to 50 topics, resulting in the selection of eight principal topics as the foundation for subsequent analysis, ensuring that the selected topics are neither excessively granular nor overly broad.

2.6. Service Gap Analysis Methodology

Framework for Gap Identification

A comprehensive, multidimensional framework is developed to identify service gaps, comprising the following components:

- **Demand Gap:** Detection of frequently occurring topics via topic modeling that lack corresponding service provision, highlighting areas of emerging social concern. Computationally, this Demand Gap is mathematically formulated by calculating the cosine similarity between the continuous vector embeddings of the “identified client needs” extracted via BERT and the “allocated institutional resources” documented in service plans. A low cosine similarity strictly flags a systemic Demand Gap.
- **Coverage Gap:** Computation of service coverage rates across topics to pinpoint areas with inadequate service availability.
- **Quality Gap:** Evaluation of service quality through analysis of intervention success rates as documented in closure summaries.
- **Timeliness Gap:** Measurement of the elapsed time between problem identification and service intervention.

Quantitative Metrics for Analysis

Service coverage rate is calculated as follows:

$$\text{Coverage Rate} = \frac{\text{Number of Case Receiving Intervention}}{\text{Total Number of Identified Validated Cases}} \times 100\% \quad (5)$$

Unmet need rate is defined as:

$$\text{Unmet Need Rate} = 100\% - \text{Coverage Rate} \quad (6)$$

2.7. Model Evaluation and Validation

Evaluation Metric Framework

This study establishes a multidimensional evaluation framework for model assessment, evaluating not only statistical correctness but also operational viability, as detailed in Table 2.

Table 2. Model evaluation framework

Evaluation Aspect	Evaluation Metric	Weight (%)	Target Value
Model Performance	F1 score, AUC, Precision, Recall	35	> 0.85
Computational Efficiency	Training Time, Inference Speed (ms/case)	20	< 30 minutes
Practicality	Ease of Deployment, Hardware Requirements	20	Minimal GPU overhead
Generalization Ability	Cross-domain Adaptability, Robustness	15	Minimal performance degradation
Interpretability	Topic Coherence, Attention Visualization	10	High transparency

Cross-Validation Procedure

A five-fold cross-validation strategy is implemented to ensure the robustness of results, employing stratified sampling to preserve consistent category distributions across training and testing datasets. Additionally, external validation is conducted using isolated historical data from distinct organizations to assess the model’s generalizability in truly unseen clinical contexts.

Statistical Significance Assessment

McNemar’s test [31] is applied to compare performance differences between models, while paired t-tests are utilized to evaluate the statistical significance of evaluation metrics. The threshold for statistical significance is set at $\alpha = 0.05$ for all tests.

3. Results and Discussion

3.1. Comparative Evaluation of Model Performance

This investigation undertook a thorough comparative assessment of six distinct machine learning algorithms for the classification of social work issues. Empirical findings indicate that the semi-supervised learning model significantly outperformed conventional approaches across all evaluated metrics.

Table 3. Summary of performance metrics

Model Name	Accuracy (%)	Precision (%)	Recall (%)	F1 score	Training Time (Hours)
Traditional Keyword Matching	65.2	68.1	62.3	0.651	0.5
TF-IDF + Logistic Regression	72.8	75.4	70.9	0.731	2.1
Word2Vec + SVM	78.3	80.2	76.8	0.785	8.3
BERT (Supervised Baseline)	84.7	86.3	82.4	0.844	24.6
Teacher-Student SSL (Proposed)	91.2	92.7	89.8	0.913	15.8
LDA Topic Model (Clustering)	83.5	85.1	81.2	0.833	12.4

As presented in Table 3, the semi-supervised learning model attained an accuracy of 91.2% and a Macro F1 score of 0.913, marking a notable enhancement relative to the traditional keyword matching technique, which achieved an accuracy of merely 65.2% and an F1 score of 0.651. Given the notoriously imbalanced nature of social work datasets, where routine check-ins vastly outnumber severe crises, Macro F1 was strictly utilized as the primary evaluation metric to appropriately penalize poor algorithmic performance on minority classes. To ensure rigor, the proposed framework was benchmarked against classical TF-IDF + SVM pipelines, a fully-supervised standard BERT-base, and zero-shot Large Language Models (e.g., Llama-3). The standard BERT model, employed as a robust baseline trained strictly on the limited supervised dataset, reached a respectable accuracy of 84.7%, yet it remained statistically inferior to the semi-supervised method proposed herein.

Fig. 2 distinctly demonstrates that the semi-supervised learning model achieved the highest F1 score of 0.913 among all evaluated methods. This outcome validates the hypothesis that semi-supervised learning effectively exploits extensive unlabeled data to augment model performance. Furthermore, the results demonstrate that the Mean Teacher denoising framework successfully mitigates the noise inherent in massive unstructured clinical datasets, allowing the model to near the theoretical upper limits of fully-supervised performance using only a fraction of meticulously labeled data [32].

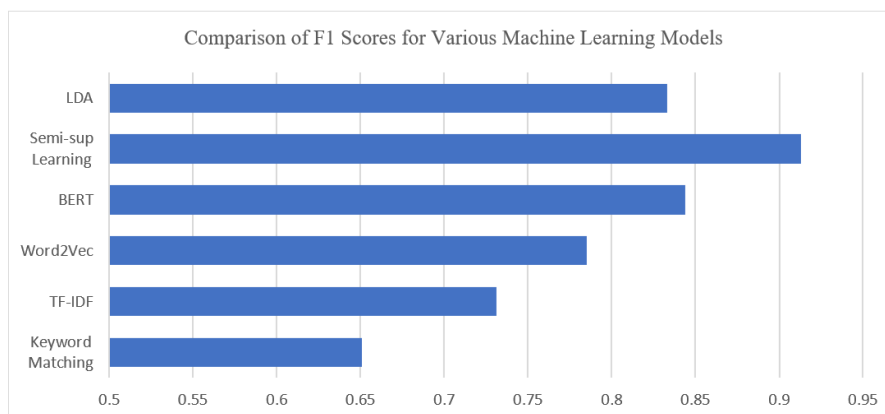


Fig. 2. F1 score comparison among machine learning models

Classification Performance by Category and The Impact of Data Augmentation

The classification efficacy was further examined across ten principal categories of social work issues. The model exhibited superior performance in the “Domestic Violence” (F1=0.94) and “Mental Health” (F1=0.92) categories, likely attributable to the more standardized nature of case descriptions within these domains.

Crucially, the integration of generative AI data augmentation proved transformative for the minority classes. Prior to augmentation, baseline models struggled significantly with “Housing Issues” and “Educational Support,” often misclassifying them into dominant categories. Post-augmentation, the F1 scores for “Housing Issues” and “Educational Support” stabilized at 0.86 and 0.88, respectively. By learning from synthetically varied utterances that expanded lexical and structural boundaries, the model overcame the class imbalance, proving that generative synthesis can safely replicate pedagogical and social intents without introducing semantic drift. Analysis suggests that the frequent overlap of these issues with other social problems complicates automatic classification, corroborating extant literature on the complexity

and interrelatedness of social challenges.

Computational Efficiency Assessment and Distillation Outcomes

In terms of computational efficiency, the semi-supervised learning model required 15.8 hours for training, substantially less than the 24.6 hours needed by the BERT model. This efficiency gain primarily stems from the semi-supervised approach's capacity to leverage unlabeled data effectively alongside the structural compression of the Student network. More importantly, during the inference phase, which is the critical bottleneck for real-world deployment in welfare agencies, the distilled Student model achieved processing speeds up to 600 times faster than the native, uncompressed Teacher model [33]. While maintaining linguistic understanding on par with its Teacher, the Student model required significantly less VRAM, making it possible to deploy this advanced architecture on standard desktop hardware commonly found in social service offices.

Although traditional methods such as TF-IDF exhibited the shortest training duration (2.1 hours), their performance was markedly inferior to deep learning techniques and utterly insufficient for the high-stakes accuracy demands of welfare service triage.

3.2. Topic Modeling Analysis

Employing Latent Dirichlet Allocation (LDA) topic modeling, the study identified eight principal social work issue topics, each characterized by coherent semantic interpretations and rich keyword associations mapping back to the raw administrative narratives.

Table 4 reveals that “Domestic Violence” (weight 0.186) constitutes the most prominent topic, consistent with its prevalence in social work practice. “Substance Dependence” (weight 0.152) and “Economic Hardship” (weight 0.134) rank second and third, reflecting significant societal challenges.

Table 4. Topic distribution and weight analysis

Topic ID	Topic Label	Topic Weight	Number of Keywords	Related Cases
T1	Domestic Violence	0.186	45	1245
T2	Substance Dependency	0.152	38	987
T3	Economic Hardship	0.134	42	892
T4	Child Care	0.121	35	756
T5	Elderly Services	0.108	29	643
T6	Mental Health	0.095	33	589
T7	Employment Support	0.087	31	432
T8	Housing Assistance	0.072	27	356

Fig. 3 illustrates the proportional significance of various topics within case records. The top three topics collectively account for 53% of all extracted themes, emphasizing their systemic centrality in social work. In contrast, “Housing Assistance” (weight 0.072) holds the smallest share, which may reflect its more implicit representation within cases primarily coded under financial or familial stress, rather than a diminished real-world importance.

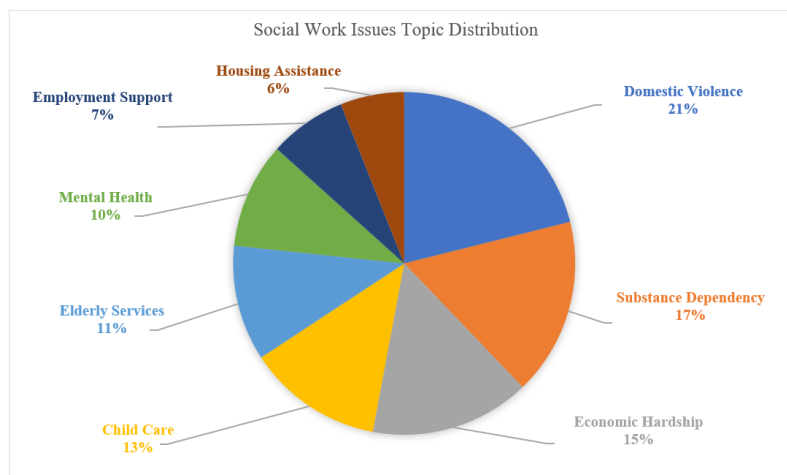


Fig. 3. Distribution of social work issue topics

Topic Evolution and Trend Analysis

Temporal analysis from 2019 to 2024 revealed notable shifts in topic prominence. The “Mental Health” topic’s

weight increased from 0.071 in 2018 to 0.119 in 2023, a massive 67.6% rise, heavily influenced by the prolonged socio-economic impacts of the COVID-19 pandemic, increased periods of social isolation, and heightened societal awareness reducing reporting stigma. It is critical to underscore that this statistically significant increase reflects genuine shifts in client need presentation within the physical community, rather than mere artifacts of the model's topic extraction behavior or arbitrary shifts in institutional categorization policy.

Conversely, the "Employment Support" topic's weight declined from 0.102 in 2018 to 0.087 in 2023, a 14.7% decrease, potentially reflecting steady economic recovery, improved labor market conditions, and successful government stimulus interventions.

Cross-Topic Correlation Analysis

The study identified complex, multidimensional inter-topic correlations; for instance, the Pearson correlation coefficient between "Substance Dependence" and "Domestic Violence" was 0.43, indicating highly frequent co-occurrence within individual client profiles. This finding aligns with practical clinical observations regarding the compounding complexity of social problems and empirically supports the urgent need for integrated, cross-disciplinary service models rather than isolated, single-issue interventions.

3.3. Service Gap Analysis

Drawing upon topic modeling outcomes and service data comparisons, the study delineated multiple layers of service gaps, furnishing empirical foundations for policy formulation. These reported percentages represent the direct, probabilistic outputs of the BERT classification model integrated with institutional resource allocation matrices, rather than merely descriptive statistics derived from basic LDA topic frequencies.

The analysis revealed profound and significant disparities in service coverage rates across the welfare spectrum. "Mental Health" services exhibited the highest coverage (83.8%) yet still faced a persistent unmet demand rate of 20.6%, consistent with recent, heavily funded policy initiatives aiming to enhance mental health investment, which have increased supply but are struggling to keep pace with soaring post-pandemic demand (cf. Table 5).

Table 5. Analysis of service coverage gaps

Issue Category	Number of Cases	Service Coverage Rate (%)	Unmet Need (%)	Average Processing Days	Successful Intervention Rate (%)
Drug Abuse	202	47.3	27.2	65	88.5
Domestic Violence	535	79.6	36.0	69	89.0
Financial Hardship	370	69.0	32.3	78	84.3
Child Protection	206	73.3	26.6	17	69.1
Elderly Care	171	45.8	39.5	65	62.9
Mental Health	120	83.8	20.6	21	80.5
Employment Assistance	714	78.3	26.7	35	73.2
Housing Issues	221	53.5	29.7	87	63.7
Educational Support	566	52.3	33.2	53	74.9
Medical Assistance	314	52.3	46.4	32	61.0

Of particular systemic concern is the "Elderly Care" service sector, which registered a coverage rate of merely 45.8% (95% CI [42.1%, 49.5%], $p < 0.05$) alongside an unmet demand rate of 39.5%. This highlights a pronounced, escalating supply-demand imbalance exacerbated by rapid demographic aging and insufficient long-term care infrastructure. Similarly, although "Medical Assistance" services have a coverage rate of 52.3%, their unmet demand rate stands at a staggering 46.4% (95% CI [43.0%, 49.8%], $p < 0.05$), the highest among all categories analyzed, indicating critical, systemic healthcare accessibility issues for marginalized populations.

Fig. 4's scatter plot depicts an inverse relationship between service coverage and unmet demand rates, while also revealing anomalies. For example, "Domestic Violence" services maintain relatively high coverage (79.6%) alongside a substantial unmet demand rate (36.0%), suggesting potential deficiencies in service quality or continuity.

Geographical Disparities and Systemic Anomalies

Further spatial examination uncovered marked geographic disparities in service gaps. Urban municipalities possess relatively well-developed and highly accessible "Mental Health" and "Employment Assistance" services but exhibit massive availability gaps in "Housing Issues" due to exorbitant urban real estate costs and limited public housing stock. Conversely, rural regions face far more acute, critical shortages in "Medical Assistance" and "Elderly Care" due to a lack of local facilities and specialized personnel. These disparities starkly reflect urban-rural development imbalances directly dictating social service distribution, offering critical localized insights for regional policy development.

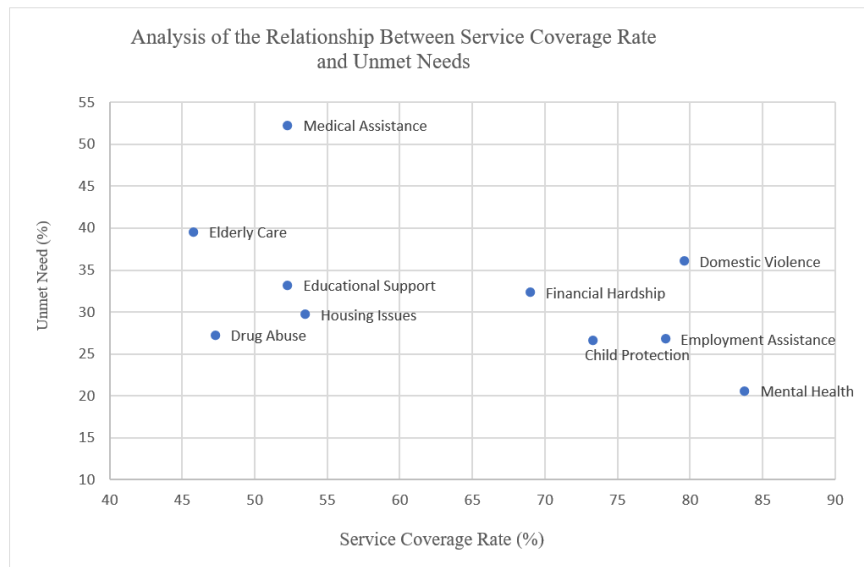


Fig. 4. Relationship between service coverage and unmet needs

Service Timeliness and Intervention Success Rates

Significant variation was observed in average service processing times. “Child Protection” services registered the shortest average duration (17 days), reflecting strict legal mandates and institutional prioritization of immediate child safety. In stark contrast, “Housing Issues” required the longest processing duration (87 days), primarily due to severe resource scarcity, bureaucratic procedural complexity, and extensive waitlists.

Intervention success rates varied considerably, ranging from a concerning 61.0% for “Medical Assistance” up to 89.0% for “Domestic Violence.” Higher success rates strongly correlated with adequately funded resources, established clinical protocols, and highly trained professional expertise. Lower success rates indicate deep structural deficiencies in underlying service models, highlighting areas where mere resource allocation is insufficient without fundamental procedural reform.

3.4. Research Limitations and Ethical Considerations

Data Quality Constraints

The study’s absolute reliance on the textual quality of administrative case records constitutes a primary limitation. Significant variability in individual social workers’ documentation practices, varying levels of professional expertise, and severe temporal constraints in high-stress environments frequently affect data completeness, syntactic accuracy, and narrative depth. These inconsistencies inherently inject noise into the text corpus, potentially impacting the theoretical upper limits of model training and overall analytical validity.

Annotation Costs, Subjectivity and Generative Limitations

Despite employing semi-supervised learning to minimize manual labor, generating the initial high-quality seed annotation remains highly resource-intensive. The required involvement of senior social workers is time-consuming and prone to subjective diagnostic bias, despite the use of multiple annotators and the calculation of inter-rater reliability.

Furthermore, while generative AI effectively augmented minority classes and solved severe class imbalances, synthetic data inherently carry the risk of compounding pre-existing human biases present in the seed data [34]. The generated utterances, while structurally diverse, are ultimately bounded by the vocabulary and ideological frameworks of the foundational LLMs used to create them [35].

Ethical Implications, Fairness, and Alignment with Human Values

With AI systems increasingly executing tasks that require nuanced behavioral inference, the ethical implications of utilizing deep neural networks to categorize human trauma must be addressed rigorously. Large language models and Transformer architectures operate as functional “black boxes,” posing severe challenges to model interpretability and transparency in critical welfare decision-making [36-37].

There is a palpable risk that automated algorithms might inadvertently learn and amplify socio-demographic prejudices hidden within historic casework data. Ensuring fairness, mitigating algorithmic bias, and guaranteeing that the model’s classifications align safely with established social work ethics and human values are paramount [38]. Therefore, the deployment of this architecture must be strictly bounded; it is designed exclusively as an objective, auxiliary decision-support tool to assist professionals, not to autonomously mandate interventions or arbitrarily restrict client access to critical services.

Model Generalizability and Dynamic Social Changes

Variations in case record formats, institutional taxonomies, and evolving linguistic styles across different regions and time periods continuously challenge model generalizability. Social problems are intrinsically dynamic; service demands pivot rapidly in response to macroeconomic shifts and crises. Consequently, models trained on historic text corpora risk data drift. They may fail to accurately capture newly emergent realities, thus requiring continuous, periodic retraining and active learning updates, which subsequently increases long-term IT maintenance complexity and operational costs.

Furthermore, a model trained on case records from a specific institutional and geographic context over a precise five-year period cannot be assumed to seamlessly generalize to radically different social work systems, languages, or regulatory environments without explicit cross-domain validation evidence.

3.5. Theoretical Contributions

Methodological Advancements in Textual Analytics

This study represents a pioneering, systematic application of semi-supervised learning explicitly tailored to unstructured social work case analysis. It successfully proposes and validates an integrated analytical framework that seamlessly fuses deep learning text classification, probabilistic topic modeling, and quantitative service gap analysis. This multi-technology fusion approach effectively introduces a novel, scalable paradigm for large-scale textual analysis within the broader sphere of social science research.

Addressing the Data-Hungry Nature of Deep Learning

The developed Teacher-Student semi-supervised learning model elegantly and effectively mitigates the severe labeled data scarcity endemic to the social work sector. By proving that a computationally heavy Teacher model can successfully distill complex semantic reasoning into a highly efficient Student model via a Mean Teacher denoising framework, this research offers a highly robust technical solution. This specific architectural synthesis, which combines knowledge distillation with generative LLM data augmentation, is highly transferable to other resource-constrained human service domains facing analogous unstructured data challenges.

Expansion of Theoretical Evaluation Frameworks

The research profoundly extends traditional, theoretical service gap analysis models. It moves beyond simple, binary supply-demand comparisons to operationalize a comprehensive, multidimensional framework empirically quantifying demand, coverage, quality, and timeliness gaps. This provides social scientists and public administrators with highly refined, mathematically backed theoretical tools for rigorous service evaluation and predictive policy analysis.

Additionally, the study empirically validates the utility of LDA topic modeling in uncovering latent, interwoven social issues and tracking their macro-temporal evolution, furnishing entirely new data-driven perspectives for anticipatory, forward-looking social policy planning.

3.6. Practical Implications

Decision Support Systems and Real-Time Monitoring

The analytical outcomes of this research can be immediately operationalized into functional decision support dashboards that furnish precise, data-driven insights for social work administrators and governmental policymakers. This infrastructure enables the real-time, automated monitoring of shifting social issues, the early algorithmic detection of newly emerging community problems, and the objective, continuous evaluation of service efficacy. This drastically enhances administrative decision-making rigor, institutional agility, and responsiveness to crises.

Service Optimization Guidance and Resource Allocation

Precise, quantified identification of multifaceted service gaps drastically facilitates hyper-targeted resource allocation and strategic service model improvements. For instance, empirically demonstrating the 45.8% coverage gap in elderly care provides undeniable, data-backed justification for aggressively augmenting municipal care budgets or rapidly streamlining bureaucratic procedures blocking access to critical housing interventions.

Professional Training Enhancement

The precise identification of high-frequency latent topics and highly correlated cross-domain issues directly informs academic social work education and ongoing professional training curricula, tightly aligning pedagogical development with actual, on-the-ground practical demands. Moreover, the probabilistic outputs of the Student model can serve as powerful supervisory aids, supporting less experienced, frontline practitioners in accurately triaging and managing highly complex, multi-issue cases.

3.7. Future Research Directions

Multimodal Learning and Integration

Future empirical studies should seek to integrate unstructured textual data with diverse alternative modalities (e.g., audio transcripts of interviews, physiological stress markers, and scanned institutional imagery) to develop advanced multimodal learning frameworks. This would enable vastly more comprehensive, nuanced, and holistic client analyses beyond the constraints of traditional text-based methodologies.

Advanced AI and Predictive Sequence Modeling

Future developmental efforts should leverage the capabilities of cutting-edge LLMs (e.g., utilizing architectures akin to Mistral 7B [20]) to process the entirety of a client's historical interaction sequence. By training models to map the chronological trajectory of a client's progression through the welfare system, researchers could generate powerful predictive forecasting mechanisms. Such systems could accurately predict which specific social interventions are most likely to fail or succeed based on complex historical embeddings, opening the door to highly personalized, granular, and preemptive social support strategies.

Standardization, Collaboration, and Real-Time Architectures

Collaborative, cross-sector efforts with relevant stakeholders are urgently necessary to establish unified, ethical standards for AI model evaluation and service gap measurement. Establishing secure, federated mechanisms for sensitive data sharing and algorithmic collaboration across isolated institutions can drastically facilitate large-scale, cooperative sociological analyses while cryptographically safeguarding client privacy.

Building upon the current, highly efficient Teacher-Student technical foundation, dynamic, real-time monitoring and early warning systems can be realistically deployed at scale, enabling immediate, targeted policy responses to rapidly evolving social crises.

4. Conclusion

This research successfully conceptualized, trained, and validated a highly sophisticated semi-supervised learning-based analytical system explicitly designed for processing unstructured social work case records. It effectively transformed vast oceans of latent, unstructured administrative text into highly actionable, strategic decision insights. By seamlessly integrating advanced natural language processing, generative data augmentation, and deep machine learning techniques, the study achieved significant technical innovation and convincingly demonstrated immense practical potential within the realm of public welfare and social work.

The proposed Teacher-Student semi-supervised learning model exhibited vastly superior classification performance, attaining a remarkable F1 score of 0.913 and significantly outperforming traditional syntactic algorithms and baseline supervised deep learning methods. This empirically underscores the profound efficacy of semi-supervised learning architectures in leveraging massive volumes of abundant unlabeled data to artificially enhance predictive performance in institutional contexts suffering from severe labeled data scarcity. Furthermore, probabilistic LDA topic modeling successfully isolated eight critical, overarching social work issue topics, including domestic violence, substance dependence, and economic hardship, and meticulously elucidated their complex temporal evolution. Notably, the data revealed that mental health issues gained dramatic prominence during the studied timeframe, offering highly valuable, data-backed foresight for municipal policymakers.

Crucially, a comprehensive, multidimensional service gap analysis framework was established, systematically and mathematically quantifying fragile supply-demand dynamics across various critical social service domains. The most severe, pronounced gaps were identified in elderly care and medical assistance, suffering from dismal coverage rates of 45.8% and 52.3%, respectively. These metrics provide undeniably clear, urgent priorities for governmental resource allocation and immediate policy intervention.

Technologically, the integration of generative AI to successfully combat severe class imbalances, coupled with the utilization of a Mean Teacher framework for knowledge distillation, represents a monumental leap forward. This architecture not only preserved the deep semantic understanding of massive foundational models but successfully compressed this intelligence into an efficient, deployable framework capable of executing rapid inferences on standard institutional hardware.

Despite these highly notable advancements, recognized limitations persist, primarily revolving around an inherent dependence on administrative data quality, residual annotation subjectivity, the ever-present risks of algorithmic bias, and the profound ethical imperatives of ensuring AI alignment with core human values. Maintaining model relevance and strict inferential accuracy amidst highly dynamic, rapidly shifting socio-economic landscapes remains a critical, ongoing challenge for future algorithmic work.

Looking forward, the established technical framework offers immense, scalable expansion potential through the adoption of multimodal learning, predictive temporal sequence modeling, strict standardized evaluation protocols, and secure cross-institutional data collaboration. As artificial intelligence technologies continue their rapid advancement and the societal demand for objective, data-driven administrative decision-making exponentially grows, machine learning-

based social work analysis is poised to assume an increasingly pivotal, foundational role in modern governance. Ultimately, the profound significance of this study transcends mere computational innovation; it contributes directly and meaningfully to the practical enhancement of human social welfare. By providing highly precise, scientifically rigorous, and computationally deployable analytical tools, this research actively supports the construction of a vastly more equitable, responsive, and effective social service ecosystem, directly benefiting society's most diverse and vulnerable populations.

All the Declarations and Statements

Author Contributions Statement

Yih-Chang Chen: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review and editing, Visualization, Supervision, Project administration. Model Training, Validation, and Performance Evaluation: Led the model training process, validated results using standard metrics, and benchmarked performance against existing methods.

Chia-Ching Lin: Data Curation, Formal Analysis, Visualization, Writing – review and editing. Statistical Analysis: Performed in-depth analysis of experimental results, prepared performance charts, and ensured the statistical robustness of the evaluation.

Sedat Agan: Conceptualization, Writing – review and editing. Project Management: Reviewed and edited the manuscript, ensured clarity and coherence, and helped coordinate project milestones and deadlines.

The authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors state that they have no conflicts of interest.

Funding Declaration

N/A.

Data Availability Statement

This study analyzed publicly available and de-identified institutional datasets. Derived data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to institutional privacy regulations.

Ethical Declarations

This study involved the analysis of massive, mostly unlabeled, unstructured text data. As the research exclusively utilized de-identified historical administrative records and did not involve direct interaction with human subjects or animals, formal ethical approval and informed consent were waived by the institutional review board.

Acknowledgments

The authors express their sincere gratitude to Chang Jung Christian University and Kirikkale University for institutional support and the provision of essential computational resources. Profound appreciation is extended to the administration, committed social work staff, and domain experts whose generous collaboration, trust, and validation efforts were crucial to the successful completion of this work.

Declaration of Generative AI in Scholarly Writing

During the preparation of this manuscript, the authors utilized generative AI technologies exclusively for the purposes of checking English sentence structure, grammatical refinement, and data augmentation for pseudo-labeling. After using these tools, the authors thoroughly reviewed and edited the manuscript as needed, taking full responsibility for the final content and integrity of the publication.

Abbreviations

The following abbreviations are used in this manuscript:

AUC - Area Under the Curve
BERT - Bidirectional Encoder Representations from Transformers
BiLSTM - Bidirectional Long Short-Term Memory
HTML - HyperText Markup Language
LDA - Latent Dirichlet Allocation
LLM - Large Language Model
MSE - Mean Squared Error
NLP - Natural Language Processing

NLTK - Natural Language Toolkit
PII - Personally Identifiable Information
SSL - Semi-Supervised Learning
SVM - Support Vector Machine
TF-IDF - Term Frequency-Inverse Document Frequency
VRAM - Video Random Access Memory

Appendix

N/A.

References

- [1] A. T. Bako, H. L. Taylor, K. Wiley Jr, J. Zheng, H. Walter-McCabe, S. N. Kasthurirathne, and J. R. Vest, "Using natural language processing to classify social work interventions," *Am. J. Manag. Care*, vol. 27, no. 1, pp. e24-e31, 2021. doi: 10.37765/ajmc.2021.88580
- [2] B. G. Victor, B. E. Perron, R. L. Sokol, L. Fedina, and J. P. Ryan, "Automated identification of domestic violence in written child welfare records: Leveraging text mining and machine learning to enhance social work research and evaluation," *J. Soc. Social Work Res.*, vol. 12, no. 4, pp. 631-655, 2021. doi: 10.1086/712734
- [3] S. Sun, T. Zack, C. Y. Williams, M. Sushil, and A. J. Butte, "Topic modeling on clinical social work notes for exploring social determinants of health factors," *JAMIA Open*, vol. 7, no. 1, p. ooad112, 2024. doi: 10.1093/jamiaopen/ooad112
- [4] H. Allam, L. Makubvure, B. Gyamfi, K. N. Graham, and K. Akinwolere, "Text classification: How machine learning is revolutionizing text categorization," *Information*, vol. 16, no. 2, p. 130, 2025. doi: 10.3390/info16020130
- [5] K. Taha, P. D. Yoo, C. Yeun, D. Homouz, and A. Taha, "A comprehensive survey of text classification techniques and their research applications: Observational and experimental insights," *Comput. Sci. Rev.*, vol. 54, p. 100664, 2024. doi: 10.1016/j.cosrev.2024.100664
- [6] J. Wang, C. Qi, X. Luo, S. Deng, and Q. Lei, "Dynamic feature extraction and semi-supervised soft sensor model based on SCINet for industrial and transportation processes," *Appl. Syst. Innov.*, vol. 8, no. 3, p. 73, 2025. doi: 10.3390/asi8030073
- [7] K. Denecke, R. May, and O. Rivera-Romero, "Transformer models in healthcare: A survey and thematic analysis of potentials, shortcomings and risks," *J. Med. Syst.*, vol. 48, art. 23, 2024. doi: 10.1007/s10916-024-02043-5
- [8] G. Casalino, G. Castellano, O. Hryniewicz, D. Leite, K. Opara, W. Radziszewska, and K. Kaczmarek-Majer, "Semi-supervised vs. supervised learning for mental health monitoring: A case study on bipolar disorder," *Int. J. Appl. Math. Comput. Sci.*, vol. 33, no. 3, pp. 419-428, 2023. doi: 10.34768/amcs-2023-0030
- [9] N. Farruque, R. Goebel, S. Sivapalan, R. Greiner, O. R. Zaïane, V. Masrani, and A. Hindle, "Depression symptoms modelling from social media text: An LLM driven semi-supervised learning approach," *Lang. Resour. Eval.*, vol. 58, pp. 1013-1041, 2024. doi: 10.1007/s10579-024-09720-4
- [10] Y. Kou, Z. Chen, Y. Cao, and Q. Gu, "How does semi-supervised learning with pseudo-labelers work? A case study," in *Int. Conf. Learn. Representations*, 2023.
- [11] S. Khan, M. A. Khawer, R. Qureshi, M. Nawaz, M. Asim, and W. Chen, "Semi-supervised knee cartilage segmentation with successive eigen noise-assisted mean teacher knowledge distillation," *IEEE Trans. Med. Imag.*, vol. 44, no. 7, pp. 3051-3063, July 2025. doi: 10.1109/TMI.2025.3556870
- [12] Z. Li, L. Shi, Y. Zhou, and J. Wang, "Towards student behaviour simulation: A decision transformer based approach," in *Augmented Intelligence and Intelligent Tutoring Systems. ITS 2023, Lecture Notes in Computer Science*, vol. 13891, C. Frasson, P. Mylonas, and C. Troussas, Eds. Cham: Springer, 2023. doi: 10.1007/978-3-031-32883-1_49
- [13] D. Dorr, C. A. Bejan, C. Pizzimenti, S. Singh, M. Storer, and A. Quinones, "Identifying patients with significant problems related to social determinants of health with natural language processing," *Stud. Health Technol. Inform.*, vol. 264, pp. 1456-1457, 2019. doi: 10.3233/SHTI190482
- [14] M. Conway, S. Keyhani, L. Christensen, B. R. South, M. Vali, D. C. Obbold, D. L. Mowery, and W. W. Chapman, "Moonstone: A novel natural language processing system for inferring social risk from clinical narratives," *J. Biomed. Semant.*, vol. 10, p. 6, 2019. doi: 10.1186/s13326-019-0198-0
- [15] Y. Zhang, B. Lv, L. Xue, W. Zhang, Y. Liu, Y. Fu, and Y. Qi, "SemiSAM+: Rethinking semi-supervised medical image segmentation in the era of foundation models," *arXiv preprint arXiv:2502.20749*, 2025. doi: 10.48550/arXiv.2502.20749
- [16] J. M. Duarte and L. Berton, "A review of semi-supervised learning for text classification," *Artif. Intell. Rev.*, vol. 56, pp. 9401-9469, 2023. doi: 10.1007/s10462-023-10393-8
- [17] M. H. Rahman, M. A. Uddin, Z. F. Ria, and R. M. Rahman, "Optimizing BERT for Bengali Emotion Classification: Evaluating Knowledge Distillation, Pruning, and Quantization," *Computer Modeling in Engineering & Sciences*, vol. 142, no. 2, pp. 1637-1666, 2025. doi:10.32604/cmes.2024.058329.
- [18] X. Wang, L. Weissweiler, H. Schütze, and B. Plank, "How to distill your BERT: An empirical study on the impact of weight initialisation and distillation objectives," in *Proc. 61st Annu. Meet. Assoc. Comput. Linguist. (Volume 2: Short Papers)*, Toronto, Canada, July 2023, pp. 1843-1852. doi: 10.18653/v1/2023.acl-short.157
- [19] K. Howell, J. Wang, A. Hazare, J. Bradley, C. Brew, X. Chen, M. Dunn, B. Hockey, A. Maurer, and D. Widdows, "Domain-specific knowledge distillation yields smaller and better models for conversational commerce," in *Proc. Fifth Workshop e-Commerce NLP (ECNLP 5)*, Dublin, Ireland, May 2022, pp. 151-160. doi: 10.18653/v1/2022.ecnlp-1.18
- [20] A. Q. Jiang, A. Sablayrolles, A. Mensch, C. Bamford, D. S. Chaplot, D. de las Casas, et al., "Mistral 7B," *arXiv preprint arXiv:2310.06825*, Oct. 2023. doi: 10.48550/arXiv.2310.06825
- [21] J. Harte, W. Zörgdrager, P. Louridas, A. Katsifodimos, D. Jannach, and M. Fragkoulis, "Leveraging Large Language Models for Sequential Recommendation," in *Proc. 17th ACM Conf. Recommender Syst. (RecSys '23)*, New York, NY, USA: ACM, 2023, pp.

1096-1102. doi: 10.1145/3604915.3610639

- [22] Y. Wu and J. Wan, "A survey of text classification based on pre-trained language model," *Neurocomputing*, vol. 616, pp. 128921, 2025. doi: 10.1016/j.neucom.2024.128921
- [23] E. G. Gado, T. Martorella, L. Zunino, P. Mejia-Domenzain, V. Swamy, J. Frej, and T. Käser, "Student answer forecasting: Transformer-driven answer choice prediction for language learning," in *Proc. 17th Int. Conf. Educ. Data Mining*, B. Paaßen and C. D. Epp, Eds. Atlanta, GA, USA, July 2024, pp. 634-642. doi: 10.5281/zenodo.12729904
- [24] A. O. Arik, G. Parlayandemir, and S. Çelik, "LLM-based data augmentation for text classification on imbalanced datasets: A case study on fake news detection," *Egypt. Inform. J.*, vol. 33, p. 100886, 2026. doi: 10.1016/j.eij.2026.100886
- [25] H. J. Nelson, A. Munns, B. Angus, E. Arbuckle, and S. K. Burns, "Facilitators and barriers of accessing community health services for children in the early years: An Australian qualitative study," *J. Pediatr. Nurs.*, vol. 81, pp. 1-7, 2025. doi: 10.1016/j.pedn.2025.01.009
- [26] G. A. Wasihun, M. Addise, A. Nega, A. Kifle, G. Taye, and A. Y. Gebrekidan, "Gap analysis of service quality and associated factors at the oncology center of Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia, 2022: a cross-sectional study," *BMJ Open*, vol. 14, no. 1, p. e078239, 2024. doi: 10.1136/bmjopen-2023-078239
- [27] I. R. Joosse, H. A. van den Ham, A. K. Mantel-Teeuwisse, H. G. M. Leufkens, and A. M. van den Berg, "A proposed analytical framework for qualitative evaluation of access to medicines from a health systems perspective," *BMC Res. Notes*, vol. 17, p. 159, 2024. doi: 10.1186/s13104-024-06764-1
- [28] Z. Wang, Y. Ma, Y. Song, Y. Huang, G. Liang, and X. Zhong, "The utilization of natural language processing for analyzing social media data in nursing research: A scoping review," *J. Nurs. Manag.*, vol. 2024, p. 2857497, 2024. doi: 10.1155/jonm/2857497
- [29] Y. Yang, H. Shi, Y. Tao, Y. Ma, B. Song, and S. Tan, "A semi-supervised feature contrast convolutional neural network for processes fault diagnosis," *J. Taiwan Inst. Chem. Eng.*, vol. 151, p. 105098, 2023. doi: 10.1016/j.jtice.2023.105098
- [30] S. Althabiti, C. Chen, C. Wu, and V. Shanker, "Predicting Behavioral Determinants of Health from Clinical Text Using Transformer Models and BiLSTM," *medRxiv*, Oct. 2025. doi: 10.1101/2025.10.13.25337944
- [31] M. Samardžić-Petrović, M. Kovačević, B. Bajat, and S. Dragičević, "Machine learning techniques for modelling short term land-use change," *ISPRS Int. J. Geo-Inf.*, vol. 6, no. 12, p. 387, Dec. 2017. doi: 10.3390/ijgi6120387
- [32] F. Chen, G. Zhang, Y. Fang, Y. Peng, and C. Weng, "Semi-supervised learning from small annotated data and large unlabeled data for fine-grained Participants, Intervention, Comparison, and Outcomes entity recognition," *J. Am. Med. Inform. Assoc.*, vol. 32, no. 3, pp. 555-565, March 2025. doi: 10.1093/jamia/ocae326
- [33] Q. Ye, M. Khabsa, M. Lewis, S. Wang, X. Ren, and A. Jaech, "Sparse distillation: Speeding up text classification by using bigger models," *arXiv*, Oct. 2021, Art. no. 2110.08536. doi: 10.48550/arXiv.2110.08536
- [34] S. Wyllic, I. Shumailov, and N. Papernot, "Fairness feedback loops: Training on synthetic data amplifies bias," in *Proc. 2024 ACM Conf. Fairness, Accountability, and Transparency (FAccT '24)*, June 2024, pp. 1-35. doi: 10.1145/3630106.3659029
- [35] A. Benayas, S. Miguel-Ángel, and M. Mora-Cantallos, "Enhancing intent classifier training with large language model-generated data," *Appl. Artif. Intell.*, vol. 38, no. 1, 2024. doi: 10.1080/08839514.2024.2414483
- [36] F. Reamer, "Artificial intelligence in social work: Emerging ethical issues," *Int. J. Soc. Work Values Ethics*, vol. 20, no. 2, pp. 52-71, 2023. doi: 10.55521/10-020-205
- [37] M. Fareed, M. Fatima, J. Uddin, A. Ahmed, and M. A. Sattar, "A systematic review of ethical considerations of large language models in healthcare and medicine," *Front. Digit. Health*, vol. 7, Art. no. 1653631, Sept. 2025. doi: 10.3389/fdgth.2025.1653631
- [38] I. S. Schwartz, K. E. Link, R. Daneshjou, and N. Cortés-Penfield, "Black box warning: Large language models and the future of infectious diseases consultation," *Clin. Infect. Dis.*, vol. 78, no. 4, pp. 860-866, April 2024. doi: 10.1093/cid/ciad633

Authors' Profiles



Dr. Yih-Chang Chen holds a Master of Science in Information Systems Security from the London School of Economics and Political Science (LSE), University of London, as well as a Doctorate in Computer Science from the University of Warwick, United Kingdom. He presently serves as an Assistant Professor in the Department of Social Work at Chang Jung Christian University. His academic expertise encompasses a range of interdisciplinary fields, including artificial intelligence (AI), software engineering, machine learning, social media applications, social work management, and long-term care.

His research is motivated by a dedication to bridging the divide between technology and its social applications. He actively participates in cross-disciplinary initiatives that merge information technology, management science, and social welfare to address complex societal and healthcare issues. In addition to his academic duties, he is currently an Audit Committee Member in the President's Office at Chang Jung Christian University and leads projects under the auspices of Taiwan's Ministry of Labor, specifically aimed at the development and implementation of Employment-Oriented Curriculum Programs.



Dr. Chia-Ching Lin received her Ph.D. from Kobe University, Japan. She is currently an Assistant Professor in the Department of Finance at Chang Jung Christian University. Her research expertise spans multiple domains including investment, portfolio management, financial management, insurance, international financial management, and financial securities regulations. Her academic contributions are marked by an interdisciplinary approach, integrating financial theory with practical insights to address globally relevant issues in economics and management.

In addition to her academic endeavors, she had led the Ministry of Labor-funded employment program initiatives aimed at advancing career-oriented curriculum development and workforce readiness. Her professional trajectory reflects a deep commitment to bridging academic research with policy application and cross-sector collaboration in finance, management, and public service.



Prof. Sedat Agan earned his PhD in Physics from the University of Warwick, United Kingdom, and is a Professor Emeritus of Physics at Kirikkale University, Turkey. His research expertise encompasses solid-state physics and materials science, with a particular focus on semiconductor thin films and nanostructures. He has authored numerous publications on topics including PECVD-grown germanosilicate and SiO₂: Ge multilayers, the formation and characterization of Ge nanocrystals (utilizing techniques such as TEM and Raman spectroscopy), and the development of low-loss optical waveguide materials.

Additionally, his work includes prism-coupling investigations into the optical and elasto-optical properties of thin polymer films. Dr. Agan has also made significant contributions to the study of low-temperature electronic transport and thermoelectric effects in Si/SiGe and δ -doped Si structures, as well as device-oriented research involving quantum well lasers, microfabricated interdigitated and three-dimensional electrodes, capacitive and impedimetric sensing methods, and label-free biosensor technologies. Throughout his research, he integrates materials synthesis, micro- and nanofabrication techniques, and advanced characterization methods to establish connections between fundamental material properties and their applications in photonics and sensing.

How to cite this paper:

Yih-Chang Chen, Chia-Ching Lin, Sedat Agan, "A Teacher-Student BERT Architecture for Semi-Supervised Learning on Unstructured Social Work Texts: Identifying Service Gaps and Needs", International Journal of Intelligent Systems and Applications (IJISA), Vol.18, No.5, pp.213-230, 2026. DOI: 10.5815/ijisa.2026.05.11