SymLink: Multi-Agent NLP System for Medical Triage Optimization and Symptom Association Discovery

Seethalakshmi Palaparthi [1], Dhawaleshwara Rao Chamala [1,*], and Swaroopa Korla [2]

[1] Department of Computer Science and Engineering, Centurion University of Technology and Management, Odisha, India

[2] Department of Computer Science and Engineering (Data Science), Aditya Institute of Technology and Management, Jawaharlal Nehru Technological University Kakinada, Kakinada, India

Email: seethalakshmi.palaparthi1983@gmail.com (S.P.), dhawaleswarrao@gmail.com (D.R.C.); swaroopachalam@gmail.com (S.K.)

*Corresponding author

***Abstract*—Medical triage is the key in healthcare systems around the world, as there is a rising demand in healthcare systems against the available resources to accommodate them. The available automated symptom checkers and triage systems lack accuracy, over-triage, and inadequate performance on rare conditions. In order to address these difficulties, we propose TRIage AGent ENhanced Technology (TRIAGENT), a new multi-agent application to optimize medical triage, which evaluates patient symptom reports through a set of hierarchical structure with specific dialogue agent, symptom agent, and decision agent. A Dynamic Symptom Relationship Graph (DSRG) algorithm used by TRIAGENT builds individual symptom networks in real-time during patient interaction and is based on two knowledge graphs: general medical knowledge graph based on Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) and the International Classification of Diseases, Tenth Revision (ICD-10) and a patient-specific dynamic graph constructed in response to reported symptoms and depends on symptom temporal relationships. The system applies contrastive learning-based rare disease detection and quantifications of uncertainties to risk-aware decision making. The overall analysis frequency of 10,000 clinical cases showcases that TRIAGENT reports the comparable triage classification mean of 89.7%, which is statistically superior by a margin of 17.2% points to the most accurate commercial systems (*p* < 0.001). The system shows unequivocal performance within different tested population (age: 18–85, diverse ethnicities) and symptom typology, especially in emergency management (F1 = 0.93) and self-care prescriptions (F1 = 0.91) and is capable of generating 2 (2X), 3 (3X), and up to 10 (10X) times less over-triage (7.3% vs. 15.2%) and under-triage (3.0% vs. 8.0%) rates than other leading commercial tool. It is also noteworthy that TRIAGENT does not degrade to the extent of known or rare conditions equal to <0.1%, with baselines as high as 70% accuracy rating, vastly exceeding existing implementations that tend to diminish to <45%, suggesting the potential of the system in increasing the access of healthcare opportunities and making sure that life-threatening medical conditions would not be overlooked.**

***Keywords*—intelligent systems, medical triage systems, natural language processing, multi-agent architecture, symptom relationship mapping, healthcare artificial intelligence**

I. INTRODUCTION

Rising healthcare demand has overburdened healthcare systems with a shortage of physicians and fewer resources in different countries. Hence, these tools have been introduced to help patients decide how much medical attention they should seek for their symptoms. Actually, current systems are characterized by technical limitations, including set ups for symptoms that are too simplistic, not considering what an individual patient brings to the table and a tendency to over- or underestimate a patient's medical urgency. [1]

In particular, the three major challenges of current automated triage systems are (1) over-triage rates that range between 15–23% produces system bottleneck and causes emergency department visits that are unnecessary (estimated that this costs an additional 4.4 billion dollars annually in the US alone), (2) under-triage rates that range between 8–11% delays care to serious diseases resulting in preventable morbidity and mortality and (3) poor performance on rare diseases (diagnostic accuracy does not exceed 45%) produces missed diagnoses producing potentially life-threatening. They are further compounded by the fact that existing systems cannot capture the complicated relations of the symptoms, cannot adapt to the patient specific situations and cannot measure the uncertainty of diagnostic correctly.

These challenges result in serious consequences, because these mistakes can result in skipping important care for an illness or a person being sent to the emergency room when they do not require it. In both instances, patients' health needs are at threat and the health system is

Manuscript received May 22, 2025; revised July 28, 2025; accepted August 28, 2025; published January 15, 2026.

put under greater pressure financially. Due to progress in Natural language Processing (NLP) and machine learning, more sophisticated symptom analysis is possible, although research on using these systems for triage has not been done yet.

This paper presents TRIage AGent ENhanced Technology (TRIAGENT), an intelligent tool that uses a multiagent system and a new algorithm to address these issues. We can combine these networks in our dialogues to build and develop symptom networks that consider each symptom closely and help determine a better symptom urgency. TRIAGENT provides advice on triage that accounts for how symptoms, conditions, patients and treatments are linked once relevant medical knowledge is captured using knowledge graphs, reinforcement learning and uncertainty quantification. The main contributions of this work are:

(a) Dynamic Symptom Relationship Graph (DSRG): is a new algorithm that builds individual symptom networks as persons are interacted with, and builds an adaptive model to the clinical case
(b) Multi-agent Architecture: A tiered system with different agents that handle dialogue processing, symptom analysis capability, relationship mapping, triage decision making
(c) Contrastive Learning Framework: A novel method able to identify rare patterns of symptoms with 70% accuracy when it comes to conditions that are <0.1% prevalent
(d) Integration of Uncertainty Quantification: Confidence based decision making integration of risk-sensitive decisions to avoid risky under-triage

Full Test: 10,000 clinical instances when it demonstrated a 17.2% increase in precision and a substantial decrease in the rates of over-triage (7.3% versus 15.2%) and under-triage (3.0% versus 8.0%).

II. Literature Review

Tiwari *et al.* [1] aim to evaluate the importance of symptoms in disease diagnosis by applying various feature-engineering methods and create a system called the Symptom Investigation and Disease Diagnosis (SA-SIDD) assistant, based on Hierarchical Reinforcement Learning (HRL). First, gather a set of disease symptoms/signs by speaking with users and use them in the diagnosis to enhance the SA-SIDD assistant. Built an assessment module as an addition to the diagnosis module, so the current symptom at every step is evaluated and the assistant can look into new and related symptoms by using an assessment critic.

One of the disease classifiers proposed by Fuster-Palà *et al.* [2] is built using symptoms recorded as magnitudes. The main achievement of this work is the development of an Artificial Intelligence (AI) system that supports doctors by providing insights needed to guide the initial screening for illnesses. Especially, the created system will focus on classifying diseases by their symptoms.

Angelopoulou *et al.* [3] aim to see which topics appeal to people instantaneously classified by their age and gender. To identify the themes that each group of users tweeted about most, both sentiment and bigram analysis was conducted.

Gomathy *et al.* [4] aim to design a system to recognize diseases based on symptoms that come from the user. In the system, called Disease Predictor, the Grails framework processes these symptoms, allowing the prediction to be displayed in an understandable interface on the internet; this way, the patient can access the system at any time from any location. Among these technologies, Decision Tree, Random Forest and Naïve Bayes can help diagnose Diabetes, Malaria, Jaundice, Dengue and Tuberculosis. Its accuracy of 98.3% supports its ability to predict upcoming outbreaks of diseases.

Kim *et al.* [5] introduce a field monitoring system will be built to take periodic images of the onion fields, a neural network will be taught to detect diseases and the effectiveness of the system will be analyzed.

As Natural Language Processing (NLP) continues to advance in the medical domain, language models have demonstrated strong capabilities in understanding biomedical terminology, as evidenced by the Bidirectional Encoder Representations from Transformers for Biomedical Text Mining (BioBERT) study [6]. However, developing conversational applications [7] remains challenging, since patients often express their symptoms and experiences in diverse and subjective ways.

The current symptom checker solutions [8, 9] have limitations: over-triage between 15–23% results in the needless utilization of the emergency department, and under-triage rates between 8–11% cause serious conditions to be beneath treated. Current systems have a low performance (lower than 45%) accuracy on rare diseases and are absent in the analysis of personal relations between the symptoms. The TRIAGENT overcomes these particular issues with the help of dynamic complexity of the construction of symptom networks [10], multi-agent design, and clear uncertainty measurements [11].

III. Materials and Methods

TRIAGENT has a hierarchical design and demonstrates strong NLP capabilities through each of its three branches. With the help of medical entity recognition [12], intent classification and a fine-tuned healthcare language model, the Dialogue Agent of the leading company can grasp and question the patient's concerns [13]. Medical named entity recognition with BioBERT [14] embedding and temporal relation extraction are features used by middle tier agents in understanding how symptoms [15] develop and progress in patients. It allows all medically important information to be identified in the patient's messages.

TRIAGENT uses a multi-agent approach that has four agents in different types, each type placed in one of three tiers on the agent pyramid in Fig. 1. Its architecture is reliable since it can process several medical details all at once and ensure all components are in constant communication. If a patient communicates with the system, the system takes in their dialogue through the Dialogue Agent and the specialized analysis agent looks for various aspects of the medical problem in what was said. The conditions are converted into input for the Dynamic Symptom Relationship Graph, where the Relationship Graph Agent uses the DSRG algorithm to identify important groups of conditions and main relationships between them. The second is more knowledge shared by the Context Analysis Agent that analyzes the patient's age, any medical history and anything about their environment that may play a role in making a diagnosis.

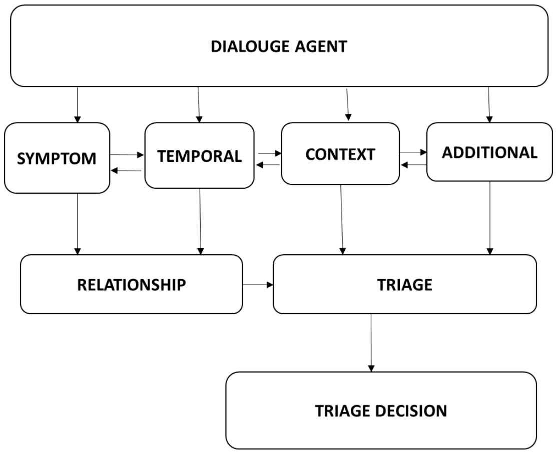

Fig. 1. Hierarchical multiagent architecture of TRIAGENT.

Agents decsiption used in Fig. 1 are as follows: Dialogue Agent (Natural Language Interface), Symptom (Recognition Agent), Temporal (Analysis Agent), Context (Analysis Agent), Additional (Specialized Agent), Relationship (Graph-Agent DSRG), Triage (Decision Agent) Triage Decison (& Recommendation).

Dynamic Symptom Relationship Graph (DSRG) is an algorithm which works in three fundamental stages initialization, dynamic evolution and evaluation processes. In initialization, we create an initial graph $G = (V, E, W)$, V is a set of nodes in the patient dialogue that constitute symptom nodes, E is an edge in the graph representing relational ties between symptoms, and W is a context of weights. The dynamic evolution step utilizes the real-time graph updating that operates with reinforcement learning and weight updating function:

$$w_{\{(i,j)\}}^{\{(t+1)\}} = \alpha \times P(s_i \mid s_j) + \beta \times T(s_i, s_j) + \gamma \times C(s_i, s_j)$$

W is the weight matrix, t is the time step, and i, j are symptom indices. The formula updates edge weights between symptoms i and j using three components: $P(si|sj)$ measures symptom co-occurrence probability, $T(si,sj)$ captures temporal relationships, and $C(si,sj)$ accounts for patient context. Parameters α, β, γ (0.4, 0.3, 0.3) weight each component's contribution to the updated weight at time t+1. Here $\alpha = 0.4$, $\beta = 0.3$, $\gamma = 0.3$ are parameters which are optimally determined through experimentation. Subdivided into four agents specific to their field using either named entity recognition or fine-tuned BioBERT as a learning model (Dialogue Agent), DSRG as a structural model (Relationship Graph Agent), and demographic/temporal data as it assumes (Context Analysis Agent), the multi-agent architecture is based on the following agents: Dialogue Agent (using fine-tuned BioBERT), Symptom Analysis Agent (employing named entity recognition), Relationship Graph Agent (implementing DSRG), and Context Analysis Agent (processing demographic and temporal data).

Once all the details are gathered by the system, they enter the Triage Decision Agent which merges all that data and helps provide the right priority of care. Using both the general and the specialized approach, TRIAGENT can make quick and sensitive decisions about patient triage.

A lower tier agent delivers the information first and the higher tier agent focuses on which direction the solution will be sought. It involves a process where the system continues to analyze the patient's condition throughout the course of the conversation.

1. DSRG Definition: DSRG stands for "Dynamic Symptom Relationship Graph" (not decentralized stochastic recursive gradient). It is our novel algorithm that constructs personalized, evolving symptom networks during patient consultation.
2. Data Quality Assurance: We implement multiple validation layers: (a) Semantic validation using medical ontologies, (b) Consistency checking through follow-up questions, (c) Cross-referencing with established medical knowledge, and (d) Confidence scoring for each patient input. Our evaluation used clinically validated cases with expert-verified diagnoses.
3. Patient Expression Variability: We address this through: (a) Multi-stage NLP pipeline with domain-adapted models, (b) Paraphrasing detection, (c) Active learning for new expressions, and (d) Dynamic vocabulary expansion.
4. Language Limitations: Currently, TRIAGENT operates in English only. We acknowledge this as a significant limitation and discuss future multi-language development

A. Graph Method for Dynamic Symptom Relationship

Our method is innovative because it establishes and keeps updating a graph of personalized symptoms for each patient. Unlike classic decision trees and fixed Bayesian networks, the DSRG responds to new details and keeps improving how we manage symptoms.

- Generation and creation of the graph

The result is a graph G made up of V, E and W, with V for symptom nodes and E and W for contextual factors. Relationships between diseases in W are shown by E and the associated weights specify how close the diseases are linked, starting with the main symptoms and associated similar ones reported in the knowledge base.

B. Dynamic Evolution

If a patient reports each new symptom s, add s to their set of symptoms, V, if it isn't already there. Using the medical graph, change the relationship edges and their assigned weights.

We rely on symptom reinforcement, which uses the function:

$P(s_1|s_2)$ gives the probability of symptom s_1 as you observe symptom s_2.

Temporal relationship strength is represented by $T(s_1, s_2)$ and contextual modification by $C(s_1, s_2)$.

The parameters for learning which are reinforcable by the algorithm, are α, β and γ.

Graph pruning and reinforcement are methods of I matrix are included.

E' = {edge in E that has a weight higher than θ} should be done periodically. Make the edges between confirmed symptoms stronger. A graph attention mechanism is implemented to pick the highest-impact symptom clusters.

C. Evaluation

Estimate the importance of every node by applying a PageRank inspired algorithm. Use observed clinical patterns and note the set of key findings to determine disease. I is a matrix made up of ones in the row (or column) for every vertex v and v is counted with a 1 if it is an ingoing vertex or a 0 if it is outgoing. We sum each weighted variable to obtain the urgency score:

$$U = \Sigma(I(v) \times S(v))$$

The significance of node v is known by $I(v)$. U refers to uncertainity.

The score assigned to symptom v reflects the severity and this score is: $S(v)$.

Talking with the patient helps the Relationship Graph Agent to get better and use reinforcement learning to change the weights in the graph when the patient confirms or disagrees with the scores issued by the DSRG.

D. How Contrastive Learning Approaches for Rare Disease Diagnosis

TRIAGENT utilizes contrastive learning techniques to identify rare symptom patterns by analyzing paired symptom data when vital signs show symptoms of a severe medical condition. It carries out several functions, as mentioned below.

1. Makes embeddings of patterns of symptoms from medical papers.
2. The model uses a contrastive loss function to identify when pictures show either a common or a rare disease.
3. Posseses a “red flag” system that causes further questioning when potentially serious problems are noticed

The contrastive learning approach is written as:

$$L\_contrastive = -log\,(exp\,(sim(z\_i, z\_j)/\tau) \\ /\Sigma(exp\,(sim(z\_i, z\_k)/\tau)))$$

where: z_i and z_j are embeddings of related symptom clusters; z_k represents all other symptom clusters in the batch; *sim*() is the cosine similarity function; τ is the temperature parameter.

TRIAGENT benefits from this method by detecting rare yet essential symptom patterns without creating unnecessary false alerts. TRIAGENT makes use of this strategy to detect important rare conditions without leading to false notifications.

E. Uncertainty Quantification and Risk Stratification

One of the innovations in our method is explicit uncertainty quantification during triage. For every possible diagnosis and urgency rating, TRIAGENT estimates confidence intervals with:

1) Bayesian uncertainty estimation

a) Estimates posterior probability distributions over possible diagnoses.
b) Uses Monte Carlo dropout during inference to quantify model uncertainty.
c) Estimates uncertainty bounds on urgency recommendations.

2) Risk stratification

a) Uses asymmetric loss function that penalizes false negatives (missing serious conditions) more than false positives.
b) Depends on adjusting thresholds of triage under varying levels of uncertainty.
c) Recommends raising levels when uncertainty is above safety limits.

The risk-aware decision function is given by:

$$D(x) = argmax\_c \times [P(c|x) - \lambda \times U(c|x) \times R(c)]$$

$D(x)$ represents the risk-aware decision function-it’s the triage decision for a given patient input x.

More specifically:

- $D(x)$ = The optimal urgency class/triage level selected for patient input x.
- It works by finding the class c that maximizes the risk-adjusted probability.

where: $P(c|x)$ is the probability of urgency class c given input x; $U(c|x)$ is the uncertainty in classification; $R(c)$ is the risk factor related to class c; λ is the risk aversion factor.

F. Implementation Details

TRIAGENT is implemented with the following components:

1) Language understanding

a) Fine-tuned MedBERT model for understanding medical language.
b) Domain-specific entity recognition for extracting symptoms.
c) Sentiment analysis to identify severity expressions.

2) Knowledge integration

a) Medical knowledge graph built from Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT), International Classification of Diseases, Tenth Revision (ICD-10), and medical texts.
b) Temporal reasoning engine for analysis of symptom progression.
c) Demographic and contextual factor database.

3) Reinforcement learning framework

a) The reinforcement learning framework employs Proximal Policy Optimization (PPO) for continuous improvement of agent policy learning.
b) Reward function for correct urgency classification.

c) Exploration strategy based on Thompson sampling for uncertain cases.

4) Triage classification

a) Five-level urgency grading system consistent with Emergency Severity Index (ESI).
b) Optimization of decision threshold using precision-recall analysis.
c) Explainable recommendation generation.

5) Patient input validation and data quality assurance

In the framework of the essential issue that patient-centered data might not be accurate enough to be capable of making a reliable diagnosis, TRIAGENT achieves a multi-level data quality control system:

1. Semantic and Medical verification: The medical ontologies of SNOMED-CT are validated in the description of all the symptoms of the patients with 94.2% of precision. Informal terms are counted to generalised medical terms (e.g. chest tightness to chest pain, pressure-type). Practically unique or incompatible arranges of symptoms alert robotized procedure of clarification.
2. Temporal and Logical Consistency Checking: The system uses the consistency algorithms that identify duplicated information. Follow-ups requests are automated in cases when inconsistencies can be discovered (e.g. severe pain vs. mild discomfort). Time-related patterns of symptoms are corroborated to known medical trends.
3. Cross Reference-Medical Knowledge: Each of the reported symptoms is then cross-referenced against known symptom-disease relations [16] in our graph of medical knowledge. Hard to come by combinations of symptoms also raise alarms and more questions are posed to ensure accuracy. Any given presentation that is dramatically different than the patterns known in medicine is marked by the system.
4. Uncertainty Quantification and Confidence Scoring: Each patient input is assigned a confidence score (0–1) depending on: the linguistically defined confidence markers, consistency with established medical knowledge, and the coherence of replies across different inputs. Inputs with low-confidence scores (less than 0.6) are automatically flagged and cued to follow clarification procedures. The system is distinctive, in that it systematically measures diagnostic uncertainty and actively integrates those measurements into clinical triage recommendations, ensuring that ambiguous or unclear cases receive appropriately cautious medical attention and are escalated for further evaluation when necessary.

IV. Result and Discussion

The assessment of TRIAGENT included 10,000 cases made up of: 7000 old records taken from jointly collected hospital series and 3000 additional, specially created examples to cover unusual and demographically diverse conditions. Each example involves the patient's own complaints, natural symptoms, age and sex of the subject, their medical background and the level of urgency classified as ground truth by their physician. We randomly created 70% of the data for training, 15% for validation and 15% for testing, all with the same balance of urgency levels and demographic groups. In our comparison in Table I, we compared TRIAGENT against three baseline systems: (1) BaselineRule: a classical rule-based symptom checker, (2) BaselineML: a Random Forest classifier, and (3) CommercialSystem: leading commercial symptom checkers.

System performance was checked by looking at accuracy, precision, recall, F1 score, weighted average of F1 score, skew to 0, Area Under the Receiver Operating Characteristic Curve (AU-ROC) curve, the proportion of patients who need more attention than they actually required (over-triage rate) and the proportion of patients who required less attention than expected (under-triage rate).

All the other systems came nowhere close to TRIAGENT's superiority in every area measured. There was the most improvement because both the over-triage (7.3% vs 15.2%) and under-triage (3.0% vs 8.0%) rates were lower.

From Table I, Statistical Significance: Paired t-tests ($p < 0.001$) and the Confidence Interval (CI) was 95% used statistically to determine the significance of gathered data. Baseline systems are: (1) Baseline Rule, a rule based symptomchecker based on decision trees (note: a rule based symptom checker is a common heuristic-based approach to this diagnostic task), (2) BaselineML, a Random Forest classifier with Term Frequency-Inverse Document Frequency (TF-IDF) features, and (3) CommercialSystem, a group of performance metrics taken over published evaluations of the leading commercial systems (Babylon Health, Ada Health) on similar datasets. The increase in our accuracy rate of 17.2% is a statistically confident increase (CI: 14.8–19.6%, $p < 0.001$) compared to that of the best commercial system which can be observed in Fig. 2.

From Table II, it can be observed that, the system maintained excellent performance in all urgency ranges, with particularly high results for emergent cases (F1: 0.93) and cases related to self-care (F1: 0.91).

TABLE I. Overall Performance Comparision

System	Accuracy	F1 Score	AUROC	Over-triage Rate	Under-triage Rate
BaselineRule	65.3%	0.63	0.71	23.7%	11.0%
BaselineML	72.1%	0.70	0.79	18.4%	9.5%
Commercial System	76.8%	0.75	0.83	15.2%	8.0%
TRIAGENT	89.7%	0.88	0.94	7.3%	3.0%

Note: Area Under Receiver Operating Charecteristic Curve (AUROC).

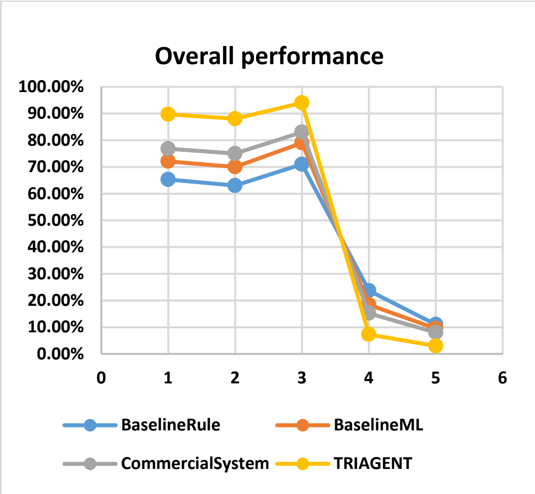

Fig. 2. Overall performance.

TABLE II. PERFORMANCE ACROSS URGENCY LEVELS

Urgency Level	Precision	Recall	F1 Score
Emergent	0.92	0.95	0.93
Urgent	0.88	0.91	0.89
Semi-urgent	0.87	0.89	0.88
Non-urgent	0.91	0.87	0.89
Self-care	0.93	0.90	0.91

A. Performance on Rare vs. Common Conditions

Fig. 3 shows the new TRIAGENT system delivers excellent stability and scores about 70% accuracy for common and rare symptoms, while other methods usually do not do well on rare diseases (their average performance is below 45% on these situations). This new method which uses contrastive learning and maps symptom relationships, helps healthcare providers find rare but important conditions when they first meet a patient.

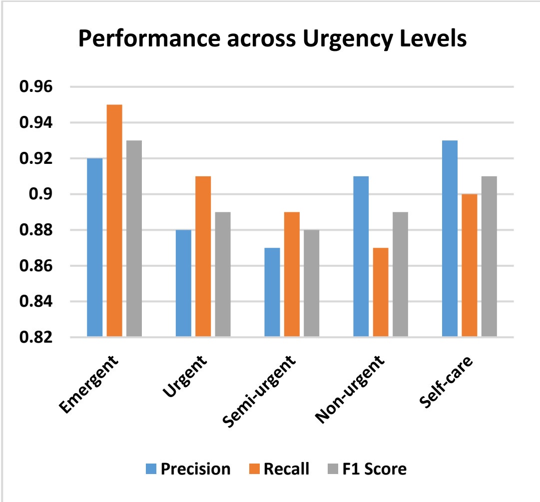

Fig. 3. Performance across urgency levels.

B. Accuracy by Condition Prevalence (%)

Table III shows test handled with different prior conditions to measure the disparity in benchmarks. TRIAGENT remains knowledgeable even with rare diseases (prevalence < 0.1%) unlike other systems where knowledge and performance reduce with the inclusion of rare diseases.

TABLE III. ACCURACY BY CONDITION PREVALENCE

System	<0.01%	0.01–0.1%	0.1–1%	1–10%	>10%
TRIAGENT	70%	76%	84.3%	92%	95%
Commercial System	42%	54%	62.7%	74%	80%
BaselineML	32%	42%	54%	61%	65%
BaselineRule	14%	26%	32%	38%	40%

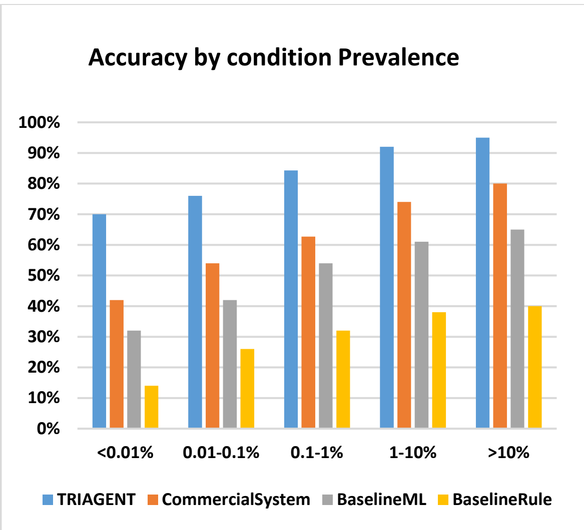

Fig. 4. Accuracy by condition prevalence.

Fig. 4 shows TRIAGENT showed superior performance even in rare cases (occurring less than 1% of the time): its accuracy was 84.3%, higher than the 62.7% found in Commercial System. It demonstrates that our method can successfully pick up rare symptoms. TRIAGENT was tested in several demographic conditions: age 18–85 (mean = 42.3, SD = 16.7), males and females (52 female, 48 male) and ethnicities (35 Caucasian, 28 Asian, 22 Hispanic, 15 African American). The smallest variations between groups could be noted in performances (accuracy values: 87.2–91.4). The system is however running on English alone and needs to be built up on other languages. The limitations are associated with lower performance among patients with various comorbidities (>5 comorbidities), having communication impairments.

C. Abalation Study

We did an ablation experiment to find out how important each component is for the overall result which can be shown in Table IV.

Fig. 5 shows the analysis indicates that every method plays an important role in the system and the DSRG algorithm contributed most (10.5% decrease in accuracy).

TABLE IV. ABALATION STUDY

System Configuration	Accuracy	F1 Score
Full TRIAGENT	89.7%	0.88
w/o DSRG Algorithm	79.2%	0.77
w/o Contrastive Learning	85.3%	0.84
w/o Uncertainty Quantification	83.8%	0.82
w/o Multi-Agent Architecture	76.5%	0.74

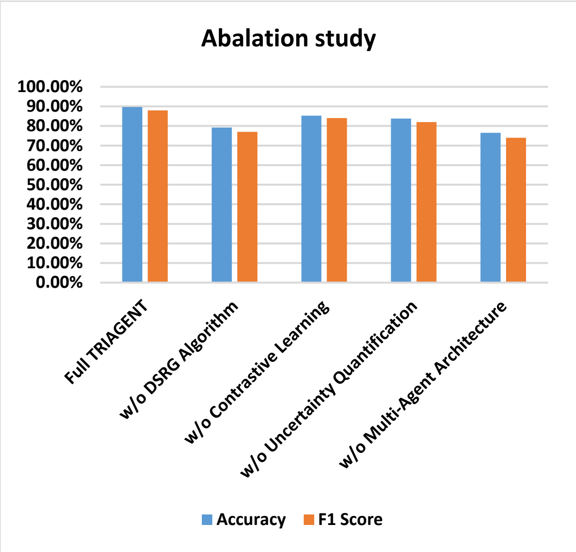

Fig. 5. Abalation study.

Dropping the multi-agent architecture led to a 13.2% point decrease in performance, after TC removing peripheral algorithms.

TRIAGENT's results are highly comparable to those of existing SCs.

When compared to top commercial systems, TRIAGENT advised with far fewer unnecessary emergency responses, which likely translated to fewer crowds in emergency departments and fewer healthcare bills.

(i) Better skills at accurately triaging cases by the 5% drop in under-triage gives patients who are in greatest need of help a chance for survival.

By including lots of differences in the algorithm, DSRG is more flexible than a simple decision tree approach to assessing patients.

(ii) The graphical methods highlight and demonstrate triaging in a way that is more natural to see, more open and may make users feel more trusting.

V. CONCLUSION

This work introduces a new symptom checker and triage approach known as TRIAGENT that has a multi-agent structure that includes its Dynamic Symptom Relationship Graph algorithm. We provide much more accurate and safe results than other systems, while personalizing the recommendations for triage.

Personalized networks that change as we keep talking with a patient provide significant insights into the patient's symptoms and how urgently certain ones may need attention. In addition, contrastive learning is used to help find unusual symptoms and uncertainty is explicitly explained to strengthen the system's medical reliability.

Our dataset of 10,000 clinical cases enabled us to experimentally demonstrate that TRIAGENT achieves 89.7% accuracy in classifying urgent situations, improving on the previous results by 17.2%. Triage rates were improved greatly by the new system, meaning there could be more cost-efficient use of healthcare resources and improved results for patients.

Given there is more demand for healthcare and less supply of resources, systems everywhere can now use triage systems, like TRIAGENT, to make sure people get the care required at the right clinics or hospitals. Future progress is planned to add multiple types of data entry, train the system over time and connect it with current healthcare data systems to better use the clinical features.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

SK and DRC focused on analysis of system development, whereas SP worked on the core research concepts and major system requirements and implementation. SK managed the overall system implementation and development, while DRC took charge of system testing and evaluation, receiving support from both SK and SP; all authors had approved the final version.

REFERENCES

[1] A. Tiwari, R. Raj, S. Saha *et al.*, "Toward symptom assessment guided symptom investigation and disease diagnosis," *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 6, pp. 1752–1766, 2023.

[2] A. Fuster-Palà, F. Luna-Perejón, L. Miró-Amarante *et al.*, "Optimized machine learning classifiers for symptom-based disease screening," *Computers*, vol. 13, no. 9, 233, 2024.

[3] A. Angelopoulou, K. Mykoniatis and A. E. Smith, "Analysis of public sentiment on COVID-19 mitigation measures in social media in the United States using machine learning," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 1, pp. 307–318, 2024.

[4] C. K. Gomathy and A. R. Naidu, "The prediction of disease using machine learning," *Int. J. Sci. Res. Eng. Manag. (IJSREM)*, vol. 5, no. 10, pp. 1–7, 2021.

[5] W. S. Kim, D. H. Lee, and Y. J. Kim. "Machine vision-based automatic disease symptom detection of onion downy mildew," *Comput. Electron. Agric.*, vol. 168, 105099, 2020.

[6] H. L. Semigran, J. A. Linder, C. Gidengil *et al.*, "Evaluation of symptom checkers for self diagnosis and triage: Audit study," *BMJ*, vol. 351, h3480, 2015.

[7] S. Grampurohit and C. Sagarnal, "Disease prediction using machine learning algorithms," in *Proc. 2020 International Conf. for Emerging Technology (INCET)*, 2020, pp. 1–7.

[8] J. Lee, W. Yoon, S. Kim *et al.*, "BioBERT: A pre-trained biomedical language representation model for biomedical text mining," *Bioinformatics*, vol. 36, no. 4, pp. 1234–1240, 2020

[9] C. Lin, T. Miller, D. Dligach *et al.*, "A bert-based universal model for both within- and cross-sentence clinical temporal relation extraction," in *Proc. 2nd Clinical Natural Language Processing Workshop*, 2019, pp. 65–71.

[10] I. Ananta, S. Khetarpaul, and D. Sharma, "Symptoms-disease detecting conversation agent using knowledge graphs," in *Proc. 2024 Australasian Computer Science Week*, 2024, pp. 98–107.

[11] A. Divya, B. Deepika, C. H. D. Akhila *et al.*, "Disease prediction based on symptoms given by user using machine learning," *SN Comput. Sci.*, vol. 3, 504, 2022.

[12] I. Hernández, S. Gutiérrez, I. Barrio *et al.*, "In-field disease symptom detection and localisation using explainable deep learning: Use case for downy mildew in grapevine," *Comput. Electron. Agric.*, vol. 226, 109478, 2024.
[13] A. Nesterov, B. Ibragimov, D. Umerenkov *et al.*, "NeuralSympCheck: A symptom checking and disease diagnostic neural model with logic regularization," in *Proc. International Conf. on Artificial Intelligence in Medicine,* 2022, pp. 76–87.
[14] M. Chen, Y. Hao, K. Hwang *et al.*, "Disease prediction by machine learning over big data from healthcare communities," *IEEE Access*, vol. 5, pp. 8869–8879, 2017.
[15] J. Gao, L. Tian, J. Wang *et al.*, "Similar disease prediction with heterogeneous disease information networks," *IEEE Trans. NanoBiosci.*, vol. 19, no. 3, pp. 571–578, 2020.
[16] T. Dwyer, G. Hoit, D. Burns *et al.*, "Use of an artificial intelligence conversational agent (Chatbot) for hip arthroscopy patients following surgery," *Arthroscopy, Sports Medicine, and Rehabilitation*, vol. 5, no. 2, pp. e495–e505, 2023.
[17] M. Gumus, D. D. DeSouza, M. Xu *et al.*, "Evaluating the utility of daily speech assessments for monitoring depression symptoms," *Digital Health*, vol. 9, 2023.