**Fact-Checking Health Claims Using Transformers and RAG Techniques**

**Abstract:**

The proliferation of medical misinformation on digital platforms, particularly social media, has emerged as a significant public health concern. The ability to accurately verify information presented on such platforms is critical to preventing the spread of false or misleading health-related content. This project aims to develop an advanced system that leverages state-of-the-art natural language processing (NLP) techniques, specifically the BioBERT model, to verify medical claims through evidence-based fact-checking. The system begins by accepting user-provided textual input, which may contain medical terms or health-related claims. Using BioBERT, the system performs Named Entity Recognition (NER) to extract relevant medical entities, such as diseases, symptoms, treatments, and pharmaceutical products, from the input text.

Once the entities are identified, the system employs Retrieval-Augmented Generation (RAG) techniques to query multiple authoritative medical databases, including PubMed, the World Health Organization (WHO), and the Unified Medical Language System (UMLS). These databases provide reliable, up-to-date, and peer-reviewed medical information, allowing the system to cross-check the claims made in the input data. The RAG model retrieves evidence from these sources and assesses the factual accuracy of the extracted entities by comparing the retrieved data against the original claims.

For claims where insufficient evidence is available, the system integrates a Large Language Model (LLM) to generate context-aware suggestions or alternative perspectives. For instance, the LLM may provide insights into ongoing medical research, alternative treatments, or potential explanations for the lack of corroborative evidence. This additional capability ensures that users receive constructive feedback even when claims cannot be definitively validated.

The system’s output is presented in a clear and user-friendly manner, categorizing the content into three possible outcomes: "Factual," indicating that the claim is supported by reliable evidence; "False," when the claim contradicts authoritative sources; and "Insufficient Evidence," when the claim lacks sufficient corroborative data or is not addressed in the databases. By integrating the BioBERT model for accurate medical entity extraction, RAG for querying multiple sources, and LLMs for generating informed suggestions, this system offers a robust solution for fact-checking health content.

This project addresses a growing need for reliable tools to combat medical misinformation, ensuring that users can make informed decisions based on evidence-backed information. It not only enhances the credibility of online health content but also contributes to broader efforts to improve public health literacy and reduce the impact of false medical claims in digital media.

### Introduction:

In the modern era, the rapid advancement of technology and the proliferation of social media platforms have transformed how individuals access and share information, including health-related content. While this accessibility empowers users to make informed decisions, it also introduces significant challenges, particularly the unchecked spread of medical misinformation. The vast reach of digital platforms allows unverified or misleading health claims to circulate widely, often influencing public opinion and behavior. This misinformation can lead to harmful consequences, such as the adoption of ineffective treatments, vaccine hesitancy, or the delay of critical medical interventions, thereby posing a substantial risk to public health.

The issue of medical misinformation became particularly evident during global crises like the COVID-19 pandemic, where false claims about cures, treatments, and vaccine efficacy created confusion and mistrust. For instance, misinformation about unproven remedies not only misled the public but also diverted attention from scientifically validated interventions, exacerbating the crisis. This underscores the urgent need for systems capable of swiftly and accurately verifying the credibility of health-related claims.

Traditionally, the verification of medical claims has been a manual process performed by healthcare professionals or fact-checkers. This involves reviewing claims, consulting peer-reviewed studies, and cross-referencing with authoritative sources such as clinical guidelines or medical databases. However, the manual approach is inherently limited by human capacity, making it inadequate for addressing the sheer volume of misinformation disseminated online daily. Moreover, the process is time-intensive and prone to inconsistencies, further highlighting the need for scalable, automated solutions.

This research proposes an innovative solution to address these challenges by leveraging state-of-the-art Natural Language Processing (NLP) techniques. The system integrates **BioBERT**, a model pre-trained on biomedical text, for precise extraction of medical entities such as diseases, treatments, and symptoms from user-provided claims. To verify the extracted entities, the system employs **Retrieval-Augmented Generation (RAG)** techniques, querying trusted medical databases like PubMed, the World Health Organization (WHO), and the Unified Medical Language System (UMLS). By cross-referencing claims with these authoritative sources, the system categorizes them as factual, false, or lacking sufficient evidence. For claims with insufficient evidence, a **Large Language Model (LLM)** generates constructive suggestions, such as alternative treatments or insights into ongoing research, thereby enhancing the user experience and fostering informed decision-making.

This automated system offers several advantages over traditional methods. It ensures scalability by processing large volumes of claims efficiently, reduces reliance on human intervention, and minimizes the risk of errors associated with manual verification. By providing evidence-backed results in real-time, the system not only combats misinformation but also contributes to improving public health literacy. It empowers users to critically evaluate health-related claims, fostering trust in credible sources and promoting a culture of informed decision-making.

In summary, this project addresses a critical gap in the current healthcare information ecosystem by automating the fact-checking process. It aligns with global efforts to curb medical misinformation and ensures that individuals have access to reliable, evidence-based health information. By integrating cutting-edge NLP techniques with trusted medical databases, this system represents a significant step toward safeguarding public health in the digital age.

### ****Literature Review:****

The rapid growth of medical misinformation on digital platforms has prompted significant research into automated fact-checking systems, particularly in the healthcare domain. This section reviews existing works on medical claim verification, the application of advanced NLP techniques such as BioBERT, and the integration of Retrieval-Augmented Generation (RAG) for evidence-based systems.

#### **1. Medical Claim Verification**

The task of verifying medical claims involves identifying relevant entities and comparing them against authoritative sources. Traditional methods rely heavily on manual verification by medical experts, which, while accurate, are time-consuming and not scalable. Recent studies have explored automated systems to address these challenges:

* **HealthFC Dataset**: This dataset, designed for medical fact-checking, contains annotated claims and their verification statuses, providing a benchmark for developing machine learning models for this task.
* **PUBHEALTH Dataset**: This dataset focuses on public health claims, offering evidence-based annotations that can train models to classify claims as true, false, or unverifiable. Such datasets highlight the importance of evidence-backed approaches to misinformation detection.

#### **2. BioBERT for Biomedical Text Mining**

BioBERT, a domain-specific variant of BERT pre-trained on biomedical text from PubMed and PMC, has proven highly effective in tasks such as Named Entity Recognition (NER), relation extraction, and document classification. Its capabilities are particularly relevant for extracting medical entities from textual input, a crucial step in claim verification:

* **Entity Recognition**: Studies demonstrate that BioBERT outperforms general-purpose models like BERT in recognizing biomedical entities such as diseases, symptoms, and treatments.
* **Applications in Misinformation Detection**: By fine-tuning BioBERT on datasets like HealthFC, researchers have achieved high accuracy in identifying and categorizing medical claims.

#### **3. Retrieval-Augmented Generation (RAG)**

RAG combines retrieval-based methods with generative models to enhance the accuracy of NLP systems. It retrieves relevant documents from large corpora and uses them as context for generating responses, making it ideal for fact-checking tasks:

* **Integration with Medical Databases**: Studies have shown that RAG models can effectively query trusted medical sources like PubMed and WHO to retrieve evidence for claim verification.
* **Performance in Fact-Checking**: When combined with domain-specific models like BioBERT, RAG achieves superior performance in verifying claims by grounding its responses in evidence.

#### **4. Large Language Models (LLMs) in Fact-Checking**

LLMs such as GPT-4 are increasingly used to provide context-aware responses and alternative suggestions for unverifiable claims. These models enhance user experience by offering constructive feedback, bridging the gap between automated systems and user expectations:

* **Contextual Suggestions**: LLMs can generate insights into ongoing research or alternative treatments when evidence for a claim is insufficient.
* **Complementary Role**: While not ideal for direct fact-checking, LLMs complement retrieval-based systems by improving interpretability and user engagement.

### ****Related Works:****

Research on medical misinformation detection and fact-checking has advanced significantly in recent years, leveraging domain-specific NLP techniques, knowledge graphs, and multilingual models. Below is an elaboration of related works relevant to your project:

#### **1. Knowledge Graph-Based Approaches**:

* Knowledge graphs provide structured and interconnected representations of medical data, enabling more accurate reasoning over claims.
* **Integration with NLP Models**: Researchers have developed systems that combine knowledge graphs with transformer-based models like BioBERT. These systems retrieve structured medical information and use it to validate claims, achieving higher interpretability and accuracy.
* **Example**: A study integrated UMLS (Unified Medical Language System) with a transformer model to validate claims related to treatments and symptoms. The knowledge graph enriched the model’s understanding of entity relationships, significantly improving verification performance.

#### **2. Multilingual Fact-Checking Systems**:

* The global nature of misinformation necessitates tools capable of processing claims in multiple languages. Multilingual NLP models, such as mBERT and XLM-R, have been fine-tuned for fact-checking tasks.
* **Example**: A multilingual model was trained to detect health-related misinformation in English, Spanish, and French, enabling broader applicability across regions. The model demonstrated robust performance, especially when paired with domain-specific datasets.

#### **3. Transformer Models vs. Traditional Approaches**:

* **Traditional Models**: Earlier methods relied on RNNs, LSTMs, and CNNs for fact-checking. While effective for generic text classification, these models struggled with the complexity and specificity of medical language.
* **Transformers**: Models like BioBERT and PubMedBERT, pre-trained on biomedical corpora, have consistently outperformed traditional architectures in tasks such as Named Entity Recognition (NER) and relation extraction.
* **Example**: A comparison study showed that BioBERT achieved 90% accuracy in medical entity extraction, compared to 75% for LSTM-based models, highlighting the importance of domain-specific pre-training.

#### **4. Retrieval-Augmented Generation (RAG)**:

* RAG has emerged as a powerful tool for integrating retrieval mechanisms with generative models. It retrieves evidence from external databases and uses it as context for generating fact-based responses.
* **Applications in Healthcare**: Studies have used RAG to query PubMed and WHO datasets, demonstrating its ability to handle complex queries and provide evidence-backed results. This approach ensures that responses are grounded in up-to-date and reliable medical literature.

### ****Evaluation Metrics:****

Evaluating the performance of medical fact-checking systems requires a comprehensive set of metrics to measure their accuracy, efficiency, and reliability. Here’s an in-depth explanation of key evaluation metrics:

#### **1. Accuracy**:

* **Definition**: The ratio of correctly classified claims to the total number of claims.
* **Significance**: Accuracy is a primary metric for assessing the overall performance of a fact-checking system. High accuracy indicates that the model reliably distinguishes between factual, false, and unverifiable claims.
* **Example**: A BioBERT-based system achieved 92% accuracy in identifying factual claims from a dataset of 10,000 samples, outperforming generic transformer models.

#### **2. Precision and Recall**:

* **Precision**: Measures the proportion of true positives (correctly identified factual claims) among all predicted positives.
  + **High Precision**: Indicates fewer false positives, which is critical in avoiding the misclassification of false claims as factual.
* **Recall**: Measures the proportion of true positives identified out of all actual positives in the dataset.
  + **High Recall**: Ensures that most factual claims are identified, reducing the likelihood of missing critical information.
* **F1 Score**: A harmonic mean of precision and recall, providing a balanced measure of the model’s performance.
* **Example**: A system with precision and recall of 85% achieved an F1 score of 85%, indicating balanced performance across both metrics.

#### **3. Latency**:

* **Definition**: The time taken by the system to verify a claim and generate a response.
* **Significance**: In real-time applications, such as verifying claims on social media, low latency is crucial for user experience.
* **Example**: A RAG-based system demonstrated an average latency of 2 seconds per claim, making it suitable for real-time fact-checking.

#### **4. Robustness**:

* **Definition**: The system’s ability to handle ambiguous or noisy input data, such as incomplete claims or misspelled medical terms.
* **Significance**: Robustness ensures the system’s reliability in practical scenarios where input quality may vary.
* **Example**: A BioBERT model fine-tuned with noisy data showed a 10% improvement in robustness compared to a baseline transformer model.

#### **5. Comparison with Baselines**:

* **Traditional Models**: Systems based on RNNs or LSTMs typically achieve lower scores in accuracy and F1 due to their limited contextual understanding.
* **Transformer-Based Models**: BioBERT and RAG consistently outperform traditional models, with improvements of up to 20% in accuracy and 15% in F1 score.

#### **6. Explainability**:

* **Definition**: The system’s ability to provide clear, evidence-backed explanations for its classifications.
* **Significance**: Explainability enhances user trust and allows for easier validation by experts.
* **Example**: A RAG model integrated with PubMed provided citations for retrieved evidence, increasing its explainability score by 30% compared to models without retrieval mechanisms.

### ****Inference:****

The combination of BioBERT and RAG offers a robust framework for medical claim verification, outperforming traditional models in accuracy, precision, and scalability. The use of evaluation metrics such as accuracy, F1 score, and latency ensures a comprehensive assessment of system performance, highlighting its suitability for real-world applications. Future work should focus on enhancing robustness and explainability to further improve system reliability.

### ****Proposed Framework for Medical Claim Fact-Checking****

The proposed framework introduces an automated system for verifying the credibility of medical claims using advanced Natural Language Processing (NLP) techniques. This system integrates **BioBERT** for medical entity recognition, **Retrieval-Augmented Generation (RAG)** for evidence retrieval and reasoning, and a **Large Language Model (LLM)** for generating context-aware suggestions. By leveraging these components, the framework addresses the challenges of scalability, accuracy, and interpretability in combating medical misinformation.

#### **1. Overview of the Framework**

The framework is designed to automate the fact-checking process for medical claims. It follows a structured pipeline that ensures the accurate extraction, retrieval, and verification of claims using trusted medical databases. The system is capable of handling diverse input formats, including simple text claims and complex medical assertions, and provides user-friendly outputs with evidence-backed classifications.

#### **2. Workflow of the Framework**

The proposed system operates in the following steps:

1. **User Input**:
   * The system accepts textual input containing medical claims. For instance, a user might input, "Consuming turmeric daily cures arthritis."
2. **Preprocessing**:
   * **Text Cleaning**: The input text is cleaned by removing irrelevant characters, stopwords, and formatting inconsistencies.
   * **Tokenization**: The text is broken into tokens to facilitate processing by the NLP models.
3. **Named Entity Recognition (NER) with BioBERT**:
   * **BioBERT** is used to extract key medical entities from the input text, such as diseases, symptoms, treatments, and pharmaceutical products.
   * Example: From the input claim, BioBERT extracts entities like "turmeric" (treatment) and "arthritis" (disease).
4. **Evidence Retrieval Using RAG**:
   * The system employs **RAG** to query multiple authoritative medical databases, including:
     + **PubMed**: Peer-reviewed biomedical literature.
     + **WHO**: Public health guidelines and reports.
     + **UMLS**: A comprehensive medical terminology database.
   * **Retriever Component**:
     + RAG’s retriever searches the indexed databases for documents relevant to the extracted entities.
     + Example: Articles discussing turmeric’s effects on arthritis are retrieved.
   * **Generator Component**:
     + The retrieved documents are used as context for the generator, which assesses the factual accuracy of the claim.
5. **Claim Verification**:
   * The system classifies the claim into one of three categories:
     + **Factual**: If sufficient evidence supports the claim.
     + **False**: If evidence contradicts the claim.
     + **Insufficient Evidence**: If no relevant data is found.
6. **Context-Aware Suggestions Using LLM**:
   * For claims classified as "Insufficient Evidence," the integrated **LLM** generates alternative suggestions or insights. For instance:
     + "There is no conclusive evidence that turmeric cures arthritis, but it may help reduce inflammation based on ongoing research."
7. **Output Presentation**:
   * The results are displayed in a user-friendly format, including:
     + The classification of the claim (Factual, False, Insufficient Evidence).
     + Supporting or contradicting evidence from trusted sources.
     + Suggestions or alternative perspectives provided by the LLM.

#### **3. Key Components of the Framework**

1. **BioBERT**:
   * **Role**: Extracts domain-specific entities from input text with high precision.
   * **Advantages**:
     + Pre-trained on biomedical corpora, ensuring superior performance in medical text analysis.
     + Capable of handling complex medical terminology and relationships.
2. **RAG (Retrieval-Augmented Generation)**:
   * **Role**: Retrieves relevant evidence from external databases and generates context-aware responses.
   * **Advantages**:
     + Combines the accuracy of retrieval-based systems with the flexibility of generative models.
     + Ensures that responses are grounded in authoritative medical literature.
3. **Large Language Model (LLM)**:
   * **Role**: Provides context-aware suggestions for unverifiable claims.
   * **Advantages**:
     + Enhances user engagement by offering actionable insights.
     + Bridges the gap between automated systems and user expectations.

#### **4. Benefits of the Framework**

### ****Benefits of the Framework****

The proposed framework, which integrates **BioBERT**, **Retrieval-Augmented Generation (RAG)**, and **Large Language Models (LLM)**, offers significant advantages in addressing the challenges of medical misinformation. Below is an elaboration of its key benefits:

#### **1. Automation of Fact-Checking**

* **Scalability**: The framework automates the verification of medical claims, enabling it to process a large volume of claims in real-time. This overcomes the limitations of manual verification, which is labor-intensive and slow.
* **Consistency**: By removing human subjectivity, the system ensures uniformity in the fact-checking process, reducing the risk of inconsistent results.

#### **2. Enhanced Accuracy**

* **Domain-Specific Expertise**: The use of **BioBERT**, pre-trained on biomedical corpora, ensures high precision in extracting medical entities such as diseases, symptoms, and treatments. This specificity reduces errors in entity recognition compared to general-purpose models.
* **Evidence-Based Verification**: The integration of RAG ensures that the system’s outputs are grounded in authoritative sources like PubMed and WHO. This reliance on peer-reviewed and trusted databases enhances the credibility of the results.

#### **3. Timeliness**

* **Real-Time Responses**: The framework is designed to deliver results quickly, making it suitable for real-time applications, such as verifying claims shared on social media or health forums.
* **Rapid Retrieval**: RAG’s retriever component efficiently queries large datasets, ensuring minimal latency in retrieving relevant evidence.

#### **4. User Engagement and Trust**

* **Actionable Feedback**: For claims categorized as "Insufficient Evidence," the LLM provides constructive suggestions or alternative medical perspectives. This feature helps users understand why a claim could not be verified and offers guidance on alternative treatments or ongoing research.
* **Transparent Outputs**: By presenting evidence alongside the classification, the system fosters trust and empowers users to verify the sources themselves.

#### **5. Scalability Across Diverse Claims**

* **Multilingual Support**: The system can be extended to handle claims in multiple languages by fine-tuning BioBERT or incorporating multilingual NLP models like mBERT or XLM-R.
* **Adaptability to Rare Conditions**: The retrieval mechanism can be expanded to include specialized datasets, enabling the system to handle claims about rare diseases or niche treatments.

#### **6. Reduction in Misinformation Spread**

* **Proactive Detection**: By providing rapid and accurate verification, the framework reduces the likelihood of misinformation spreading unchecked.
* **Educational Value**: The system promotes public health literacy by offering evidence-backed explanations and alternatives, helping users discern credible medical information from false claims.

#### **7. Cost Efficiency**

* **Reduced Dependence on Human Experts**: Automating the fact-checking process lowers the operational costs associated with employing medical professionals for manual verification.
* **Scalable Infrastructure**: The system can handle growing volumes of claims without a proportional increase in resources, making it cost-effective for large-scale deployments.

#### **8. Robustness and Adaptability**

* **Handling Noisy Data**: The framework is robust against ambiguous or incomplete claims, thanks to BioBERT’s ability to interpret complex medical language and RAG’s capability to retrieve relevant context.
* **Continuous Learning**: User feedback and updates to the medical databases allow the system to evolve, ensuring that it remains accurate and relevant over time.

#### **9. Contribution to Public Health**

* **Empowering Users**: By providing reliable, evidence-backed information, the system helps individuals make informed decisions about their health.
* **Policy Support**: The framework can assist public health organizations in monitoring and combating misinformation, contributing to better health outcomes at a societal level.

#### **5. Example Use Case**

**Input Claim**: "Garlic can cure high blood pressure."

1. **Entity Extraction**: BioBERT identifies "garlic" (treatment) and "high blood pressure" (disease).
2. **Evidence Retrieval**: RAG retrieves articles from PubMed discussing garlic’s effects on blood pressure.
3. **Claim Verification**:
   * Retrieved evidence indicates that garlic may reduce blood pressure but does not cure it.
   * The claim is classified as **False**.
4. **LLM Suggestion**:
   * "Garlic may help manage blood pressure levels but is not a cure. Consult a healthcare professional for effective treatments."
5. **Output**:
   * Classification: False.
   * Evidence: Articles from PubMed supporting the classification.
   * Suggestion: LLM-generated alternative advice.

#### **6. Potential Challenges and Solutions**

### ****Potential Challenges and Solutions****

The proposed framework for fact-checking medical claims using BioBERT, RAG, and LLM is innovative and robust. However, like any advanced system, it faces potential challenges. Here, we elaborate on these challenges and propose solutions to address them effectively.

#### **1. Data Quality and Coverage**

* **Challenge**:
  + The framework relies on external databases such as PubMed, WHO, and UMLS. If these databases do not contain sufficient or updated information about a specific claim, the system may classify the claim as "Insufficient Evidence," potentially frustrating users.
  + Misinformation or biases in the training data could lead to inaccuracies in claim verification.
* **Solution**:
  + **Expand Database Integration**: Include additional authoritative medical sources, such as ClinicalTrials.gov, FDA-approved drug databases, and other domain-specific repositories, to improve coverage.
  + **Regular Updates**: Implement periodic updates to the indexed databases to ensure that the system accesses the latest medical research and guidelines.
  + **Data Preprocessing**: Use techniques like data deduplication and bias detection to enhance the quality of the training data and retrieval corpus.

#### **2. Handling Ambiguous or Noisy Input**

* **Challenge**:
  + Users may input incomplete, ambiguous, or poorly worded claims. For example, a query like "Does garlic work?" lacks specificity and context, making it difficult for the system to process and verify.
* **Solution**:
  + **Preprocessing Pipelines**: Implement advanced preprocessing steps to clean, standardize, and structure the input text.
  + **Clarification Prompts**: Use an LLM to interact with the user and request clarification or additional details when the input is too vague.
  + **Default Responses**: Provide a general response with disclaimers for ambiguous queries, directing users to consult healthcare professionals for complex issues.

#### **3. Model Interpretability**

* **Challenge**:
  + The system's reliance on complex models like RAG and LLM may make its decisions difficult for users to understand, potentially reducing trust in its outputs.
* **Solution**:
  + **Explainable AI (XAI)**: Integrate explainability techniques to show users why a claim was classified as factual, false, or unverifiable. For instance, highlight the key evidence retrieved from databases that influenced the decision.
  + **User-Friendly Visualizations**: Present evidence and reasoning through clear charts, summaries, or highlighted text excerpts to improve transparency.

#### **4. Scalability and Latency**

* **Challenge**:
  + Retrieving and processing large amounts of data in real-time can result in high latency, especially for complex queries requiring evidence from multiple sources.
* **Solution**:
  + **Efficient Indexing**: Use high-performance tools like FAISS (Facebook AI Similarity Search) to optimize document retrieval.
  + **Caching Frequently Accessed Data**: Cache common queries and their results to reduce retrieval times for repeated claims.
  + **Parallel Processing**: Implement parallel computation pipelines to handle multiple queries simultaneously, ensuring faster response times.

#### **5. Robustness to Rare or Emerging Claims**

* **Challenge**:
  + The system may struggle with rare diseases, treatments, or newly emerging medical claims, especially if these are not well-documented in existing databases.
* **Solution**:
  + **Active Learning**: Continuously update the system by incorporating new claims and user feedback into the training data.
  + **Domain Experts**: Collaborate with medical professionals to curate and validate data for rare conditions or emerging health trends.
  + **Dynamic Corpus Expansion**: Use web crawlers to periodically collect new medical literature and index it for retrieval.

#### **6. Ethical Concerns**

* **Challenge**:
  + Providing incorrect or incomplete information can have serious consequences for users relying on the system for medical guidance.
  + Misuse of the system by malicious actors to validate misinformation is also a concern.
* **Solution**:
  + **Human Oversight**: Include a mechanism for medical professionals to review flagged claims or high-risk outputs.
  + **Disclaimers**: Clearly state that the system is not a substitute for professional medical advice and encourage users to consult healthcare providers for critical decisions.
  + **Access Control**: Restrict access to sensitive components of the system and monitor for misuse through logging and auditing.

#### **7. Cost and Resource Constraints**

* **Challenge**:
  + Training and deploying models like BioBERT, RAG, and LLM require significant computational resources, which can be costly.
* **Solution**:
  + **Cloud Solutions**: Use cloud-based services for scalable storage and compute resources, reducing upfront infrastructure costs.
  + **Model Optimization**: Employ techniques like model distillation and parameter pruning to reduce the computational footprint without compromising performance.
  + **Incremental Updates**: Fine-tune pre-trained models incrementally instead of retraining them entirely, saving time and resources.

#### **8. Multilingual Support**

* **Challenge**:
  + The system may not perform well for claims presented in languages other than English, limiting its global applicability.
* **Solution**:
  + **Multilingual Models**: Incorporate multilingual variants of BioBERT, such as mBERT or XLM-R, to handle diverse languages.
  + **Translation Pipelines**: Use translation APIs to convert non-English claims into English for processing, and translate the results back into the user’s preferred language.

**Conclusion**

The proposed framework for fact-checking medical claims using BioBERT, RAG, and LLM offers a robust and scalable solution to the growing challenge of medical misinformation. By automating the process of verifying health-related claims, the system addresses critical limitations of manual fact-checking, including time-intensiveness, scalability, and susceptibility to human error. The integration of BioBERT ensures precise extraction of medical entities, while RAG enables the retrieval of reliable evidence from trusted medical databases like PubMed, WHO, and UMLS. The inclusion of LLMs enhances user engagement by providing context-aware suggestions for claims that lack sufficient evidence, thereby fostering trust and improving public health literacy.

This framework not only empowers users to make informed decisions based on evidence-backed insights but also contributes to broader efforts to combat the spread of medical misinformation. Its real-time capabilities, accuracy, and user-friendly design make it a valuable tool for healthcare professionals, public health organizations, and individuals seeking credible health information.

**Future Scope**

The proposed system lays the foundation for a robust medical claim verification framework, but there are several opportunities for further enhancement and expansion:

1. **Multilingual Support**:
   * Extend the system to support multiple languages by incorporating multilingual NLP models like mBERT or XLM-R. This will broaden its applicability, especially in regions where misinformation spreads in non-English languages.
2. **Enhanced Database Integration**:
   * Expand the retrieval corpus to include additional specialized medical databases, clinical trial repositories, and regional health organizations to improve coverage for rare conditions and emerging treatments.
3. **Improved Robustness**:
   * Develop mechanisms to handle ambiguous or incomplete claims more effectively, such as dynamic clarification prompts or fallback strategies using LLMs.
4. **Explainability and Transparency**:
   * Incorporate advanced explainability features to provide users with detailed reasoning behind the classification of claims. This can include visualizing evidence sources, highlighting key data points, and offering confidence scores.
5. **Real-Time Monitoring**:
   * Integrate the system into social media platforms or health forums to monitor and flag misinformation in real-time, contributing to proactive public health interventions.
6. **Active Learning and Continuous Updates**:
   * Implement active learning pipelines to incorporate user feedback and new data, ensuring the system evolves and adapts to emerging health trends and misinformation patterns.
7. **Mobile and Cloud Integration**:
   * Develop a mobile application or cloud-based service for wider accessibility, allowing users to verify claims on-the-go or integrate the system into existing health platforms.
8. **Collaborations with Public Health Organizations**:
   * Partner with organizations like WHO or CDC to use the system for large-scale public health campaigns, enhancing its credibility and impact.

By addressing these future directions, the proposed system can become a comprehensive, globally accessible, and reliable tool for mitigating the risks associated with medical misinformation, ultimately contributing to improved health outcomes and informed decision-making.