Team Magma Final Project Proposal

Mauricio Bermudez, Alvaro Montoya Ruiz, Hriday Reddy Purma

Invalid Date

Final Project Proposal

University of South Florida

Mauricio Bermudez Alvaro Montoya Ruiz Hriday Reddy Purma

Professor Timothy Smith

April 7th, 2025

  Table of Content

Problem Reinforcement: 3 Motivation: 3 Plan and Methodology: 4 Data and Feasibility: 5 Literature and Best Practices: 5

# Problem Reinforcement:

We aim to address a challenge presented by primary care providers (PCPs): the inefficiency of reviewing years of fragmented outpatient records (e.g. chronic conditions, preventative care, referrals) before appointments. Specifically, our model aims to be an Anomaly-Aware Clinical Alert System. Our project proposes a model that has the ability to detect clinically significant blood glucose anomalies in patients with diabetes. Then we utilize auto-generated contextualized alerts by a fusion of time-series analysis (RAE) and clinical note summarization (BioBERT). This means that effectively we are combining RAE with BioBERT to accomplish our tasks.

Addressing potential glucose anomalies is imperative as delays can further delay interventions of hyperglycemia-induced complications (ex. Damage to the eyes, kidneys, nerves and heart. Can also increase risk of stroke or risk of foot problems). Our model hopes to be able to tackle these challenges through focus on expanding on context as existing EHR software lacks insight on information (e.g., “Glucose = 300 mg/dL”). Information presented as is lacks the depth as to WHY a patient would have a higher blood sugar level. Our model hopes to provide information such as “patient missed insulin doses”. Our goal is that our model can accurately and consistently recognize anomalies and contextualize these anomalies with corresponding patient notes. This can be then provided to a physician to create robust treatment plans for their patients to prevent any complications in the future.

# Motivation:

Administrative burdens significantly impact healthcare, with studies showing that primary-care providers spend two hours on documentation for every hour of patient care (Arndt et al., 2017). This administrative workload of manual review of fragmented data (e.g. glucose trends + notes) contributes to provider fatigue. Current alerts implemented in EHR software that trigger off these abnormal lab values, such as blood glucose spikes, lack contextual summaries which forces providers to dig through notes to correlate these anomalies with events like missed insulin doses. These inefficiencies can delay intervention and will further exacerbate burnout.

Recent studies have shown that AI-driven solutions provide promise in reducing documentation time by about 30% while also enhancing data accessibility (Paula et al., 2025). However most existing solutions fail to perform the following: 1. Fuse time-series data with text data (e.g., linking glucose spikes to clinical notes). 2. Generate actionable explanations (e.g., “Anomaly likely due to insulin omission per note”). Our model hopes to address these gaps by providing: 1. Automating anomaly detection with RAE and note summarization with BioBERT, which cuts down on manual review time. 2. Providing contextualized alerts, reducing cognitive load during time-sensitive decisions. A broader impact of our proposed model is by focusing on diabetes management – a high volume use case – we can create a scalable template for other labs (ex. International Normalized Ratio [INR] for anticoagulation therapy). This would also easily align with HL7 FHIR ecosystem, which would allow for easy integration into Epic/Cerner workflows in the future.

Our original proposal was a broad swath of “outpatient summaries”. This new proposal will allow a more focused approach of blood sugar (glucose) anomalies which allow us to still handle a sort of summarization but while focusing on a more accessible aspect of the problem. This will also still tackle administrative pain points. We also wanted to highlight our potential fusion of RAE+BioBERT as a solution to EHR limitations. By emphasizing diabetes as our entry point into historical documentation summarization and anomaly detection, we are able to utilize the techniques learned in class (e.g., autoencoders, transformers) to work on a highly prevalent problem that has clear metrics.

# Plan and Methodology:

## Part 1 - RAE Model

Our model plan begins with setting up time-series processing. We will import data from the MIMIC-III dataset comprised of glucose readings. That data can be accessed [here](https://physionet.org/content/glucose-management-mimic/1.0.0/). This dataset is sampled hourly and also provides us with timestamps of these readings. After importing the relevant data, we normalize all relevant values (μ=0, σ=1) and segment these readings into 24-hour windows to apply to RAE. The RAE model utilized will begin as so:

from tensorflow.keras.layers import LSTM, RepeatVector  
# Encoder  
encoder = LSTM(32, activation="relu")(input\_layer)  
# Decoder  
decoder = LSTM(32, activation="relu", return\_sequences=True)(RepeatVector(24)(encoder))

Upon implementation of the RAE model, we then set up an Anomaly Threshold to flag values with reconstruction error (z-score).

## Part 2 - Clinical NLP (BioBERT)

We then import the MIMIC-III discharge summaries + notes data which is accompanied by timestamps. This data can be accessed [here](https://physionet.org/content/mimiciv/3.1/). After implementation of the data, we can look to extract notes +- 2 hours from an anomaly. This model implementation would be similar to the following code:

from transformers import pipeline  
summarizer = pipeline("summarization", model="monologg/biobert-v1.1")  
summary = summarizer(note\_text, max\_length=50)

We can train and fine-tune the BioBERT model to map embeddings to templated alerts such as “Possible insulin omission per note: ‘Patient forgot injection’”.

# Data and Feasibility:

Our model will be trained on publicly available, de-identified datasets: - MIMIC-III & MIMIC-IV: MIMIC-III Glucose time-series data for training our RAE architecture, MIMIC-IV clinical data to train BioBERT to learn the necessary diabetic terminology. It will also be utilized to align timestamps and de-identify notes. [PhysioNet](https://physionet.org/) - Synthea: We can utilize this synthetic dataset to help further train the RAE model to handle outliers. This dataset specifically will generate extreme glucose values. [Synthea](https://synthea.mitre.org/downloads) - Compliance: The dataset access via PhysioNet’s credential process is free to complete, but is required to be completed to maintain HIPPA-compliance.

By leveraging these datasets, we ensure our model is trained on diverse clinical scenarios. We will implement strict data governance and compliance protocols to align with HIPAA and GDPR regulations. Additionally, we will explore synthetic data generation techniques to augment training and enhance model generalization across varied clinical settings.

# Evaluation:

## Metrics

* RAE: Precision/recall for anomaly detection (vs. clinician-flagged events).
* BioBERT: ROUGE-L score for summary quality; clinician ratings (1-5 scale).
* End-to-end latency: <10 seconds per patient (critical for real-world use).

## Success Criteria

* Precision rate of >=90% in anomaly detecion.
* Clinicians rate >=80% of summaries as “actionable.”

# Course integration

As per week 7: Autoencoders, we will utilize regularized autoencoders (RAE) to analyze time-series data in our anomaly detection aspect. Following this, week 8 should provide us the necessary insight for Transformers, where we will discuss architecture and applications of transformers. This is how we tie in our utilization of the BioBERT architecture to our model.

Following this we will utilize explainable AI (XAI), more specifically SHAP, to provide the necessary explanations of the detected anomalies. For example, XAI will provide the insights such as “Glucose spike correlated with ‘no insulin’ in the note.” We can also utlize XAI to provide the information as to why the model made the decisions it did. Outcome ideally to appear as:

Date, Anomaly, Glucose, Summary, Explanation  
2025-03-01, Yes, 320, "Pt reported skipping insulin...", "Hyperglycemia linked to insulin omission."

Changes from our original proposal was highlighting the narrower scope of focusing on blood glucose in comparison to broad outpatient summaries. We also provided explicit definition of the RAE/BioBERT fusion. Implementing clinician-rated metrics to be able to determine success criteria is also an aspect we had missed.

# Current project plan and progress updates

## Current plan

1. Team members all complete necessary credentials and training to access both the MIMIC-III and MIMIC-IV datasets.
   1. Mauricio and Hriday have completed the necessary training, awaiting feedback from Alvaro
2. Download and import MIMIC-III and MIMIC-IV datasets outlined above.
3. Begin construction of RAE architecture to identify anomalies
4. Once class covering Transformers is completed, will begin implementation of BioBERT architecture.
5. Upon verification of RAE & BioBERT model practice metrics, fusion of architecture to begin
6. Implement XAI to outline decision making metrics and expand on why anomalies may have occurred.

## Progress updates

1. Mauricio and Hriday completed training, currently building out the RAE architecture to start testing time-series data and identifying anomalies. No fusion of RAE and BioBERT at this time.

## Challenges thus far

1. Waiting on credential access and training completion review is time-consuming. Took Mauricio about a week long to receive go-ahead to access necessary module to being training. Took another few days to have training completed and reviewed to even begin access all data. To combat this, just proper time-management needed.
2. Potential of accessing non-desensitized information. Ideally we will implement scrubadub to review and remove any potential breach of PHI.

# Literature and Best Practices:

• Arndt, B. G., et al. (2017). “Tethered to the EHR: Primary Care Physician Workload Assessment Using EHR Event Log Data and Time-Motion Observations.” Annals of Family Medicine, 15(5), 419–426.  
• Landi, I., et al. (2020). “Deep Representation Learning of Electronic Health Records to Unlock Patient Stratification at Scale.” • Paula, P. A. B. de, et al. (2025). “Improving documentation quality and patient interaction with AI: A tool for transforming medical records—an experience report.” Journal of Medical Artificial Intelligence, 8(0).