Dissertation Proposal

# **Title: Transformer-Based Models for Enhancing Payment Integrity in U.S. Healthcare**

1. Introduction

U.S. healthcare processes over 5 billion claims annually, with improper payments exceeding $100 billion. Traditional fraud detection relies on static rules or shallow statistical models, which often fail to capture the semantic and sequential complexity of real-world billing data.

Transformers—originally designed for NLP (Natural Language Processing) —offer a scalable, context-aware method to model structured claims data. This research proposes a transformer-based framework for detecting fraudulent and wasteful claim patterns, with a focus on mathematical formulation, accuracy, and interpretability.

1. Motivation

* Rule-based models lack generalization to new fraud patterns.
* Transformers can model field dependencies, such as relationships between CPT (Current Procedural Terminology) codes, service dates, and provider IDs.
* The $9B payment integrity market is increasingly adopting generative and foundation models.
* This work fills a gap by tailoring mathematically grounded transformers for structured payer claims data, without relying on EHRs (Electronic Health Records).

1. Objectives

* Adapt transformer models (e.g., RoBERTa, DistilBERT: Bidirectional Encoder Representations from Transformers) to payer claim sequences.
* Mathematically formalize anomaly detection using transformers via optimization, complexity, and information theory.
* Evaluate using claims-only datasets for fraud scenarios such as duplicate claims, upcoding, and phantom billing.
* Develop interpretable and robust detection pipelines for real-world use.

1. Methodology

Data

* CMS DE-SynPUF (Medicare Claims Synthetic Public Use Files) and synthetic payer claims.
* Structured fields: CPT/ICD (International Codes for Diseases) codes, provider IDs, amounts, dates.

Model Design

* Tokenization of claim fields into embeddings.
* Fine-tuned transformer encoders with positional and provider-aware layers.
* Pairwise and time-series modelling for identifying duplicates, abnormal utilization, or conflicting code patterns.

Mathematical Tools

* PAC-Bayes (Probably Approximately Correct) bounds for generalization guarantees.
* Eigenvalue analysis of attention weights to assess feature sensitivity.
* Perturbation testing to simulate adversarial billing strategies.

Evaluation

* ROC-AUC (Area Under the Receiver Operating Characteristic Curve), F1 (metric that balances precision and recall), and cost-savings simulations.
* Bootstrapping and cross-validation.
* False-positive reduction over XGBoost and logistic baselines.

1. Duplicate Claims Detection (Use Case)

Problem:

* Duplicate claims refer to instances where the same service is billed more than once, either due to clerical errors or intentional fraud. These may differ slightly in spelling, code order, or submission date, making exact matching unreliable.

Traditional Methods:

* Rule-based systems match exact fields: provider ID, CPT/ICD code, date of service, etc.
* Fuzzy matching or string similarity (e.g., Levenshtein distance) may help but struggles with scale and semantics.

Limitations:

* High false negatives: Legitimate duplicates slip through if the field values are altered.
* High false positives: Legit claims can be wrongly flagged due to superficial similarity.

Transformer-Based Solution

**Input Representation:**

Each claim is represented as a **sequence of tokens**, including:

* + Procedure codes (e.g., CPT, ICD-10)
  + Provider ID
  + Date of service (positionally encoded)
  + Diagnosis text (unstructured)
  + Notes or justifications (free text)

**Model Architecture:**

A fine-tuned transformer (e.g., ClinicalBERT or RoBERTa) can:

* + Encode contextual semantics of claims
  + Learn dependencies between fields (e.g., "MRI of head" + "neurology dept" + "same day" → potential duplicate)
  + Compare claims pairwise or in batches using semantic similarity scores

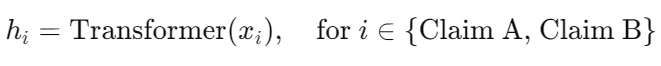
**Example:**

|  |  |
| --- | --- |
| Claim A | Claim B |
| Procedure: 99213 (Office Visit)  Provider: Dr. Smith  Date: March 10, 2025  Diagnosis: "Follow-up for hypertension"  Note: "Patient returned for medication review" | Procedure: 99213  Provider: Dr. Smith  Date: March 10, 2025  Diagnosis: "High blood pressure  "Note: "Routine follow-up, reviewed meds" |

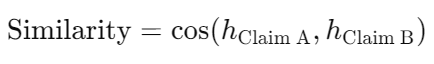
Despite minor wording differences, the transformer model will **embed both claims into the same semantic space** and identify high contextual similarity. If the model's similarity threshold (learned during training) is exceeded, the claims are flagged as potential duplicates.

**Mathematical Viewpoint:**

* Transformers convert sequences into high-dimensional vectors via:



* We then compute:



* This embedding-based similarity is far more robust than rule-based or string-matching methods, as it captures **latent meaning** and **multi-field interaction**.

**Extending to Other Scenarios**

**1. Upcoding Detection**

* **What:** Billing for a more expensive service than was provided.
* **How Transformers Help:** Learn subtle distinctions between codes (e.g., 99214 vs. 99213) and match against unstructured EHR notes or provider patterns.

**2. Medically Unnecessary Services**

* **What:** Services billed without clinical justification.
* **Transformer Advantage:** Jointly model **EHR narratives** and **claim codes** to identify misalignments (e.g., "no history of seizures" but billed for EEG: electroencephalogram).

**3. Phantom Billing**

* **What:** Billing for services never rendered.
* **Approach:** Detect anomalies in the **absence** of matching documentation or through unnatural patterns in historical provider behavior (time-series modelling).

**4. Temporal Fraud Patterns**

* Transformers with temporal position encoding can detect suspicious frequency—e.g., a provider always billing unusually consistent patient visits.

**Why Transformers Excel**

|  |  |  |
| --- | --- | --- |
| Aspect | Rule-Based | Transformer-Based |
| Handles free text | NO | YES |
| Captures semantics | NO | YES |
| Scalable to millions of claims | NO | YES |
| Learns complex, multi-field patterns | NO | YES |
| Adaptable to new fraud types | NO | YES |

Transformers provide a unified framework to handle structured and unstructured data, model relationships across fields and time, and detect patterns that are **semantically complex, morphologically varied, and contextually nuanced**—all critical in ensuring healthcare payment integrity.

1. Tools & Resources

* Ollama LLMs for lightweight model prototyping.
* OpenAI, Replicate, LangChain, and spaCy Transformers for deep model tuning.
* CMS SynPUF for publicly accessible claims data.
* Python, PyTorch, SciPy for statistical modelling.