REVA UNIVERSITY
Bengaluru, India

A Project Report on

AI-Enabled Automation Solution for Utilization Management in Healthcare Insurance

Submitted in Partial Fulfilment for Award of Degree of

Master of Technology
In Artificial Intelligence

Submitted By
Gaurav Karki
R20MTA02

Under the Guidance of
Dr. Jay Bharateesh Simha
Chief Mentor at RACE, Reva University, and CTO, ABIBA Systems

REVA Academy for Corporate Excellence - RACE
REVA University
Rukmini Knowledge Park, Kattigenahalli, Yelahanka, Bengaluru - 560 064 race.reva.edu.in

August, 2022

REVA UNIVERSITY

Bengaluru, India

Candidate's Declaration

I, **Gaurav Karki** hereby declare that I have completed the project work towards the Master of Technology in Artificial Intelligence at, REVA University on the topic entitled **"AI-Enabled Automation Solution for Utilization Management in Healthcare Insurance"** under the supervision of **Dr. Jay Bharateesh Simha.** This report embodies the original work done by me in partial fulfilment of the requirements for the award of degree for the academic year **2022.**

Place: Bengaluru

Date: 20th August 2022

Name of the Student: **Gaurav Karki**

Signature of Student

REVA UNIVERSITY
Bengaluru, India

Certificate

This is to Certify that the project work entitled **"AI-Enabled Automation Solution for Utilization Management in Healthcare Insurance"** carried out by **Gaurav Karki** with **SRN R20MTA02,** is a bonafide student of REVA University, is submitting the project report in fulfilment for the award of Master of Technology in Artificial Intelligence during the academic year **2022**. The Project report has been tested for plagiarism, and has passed the plagiarism test with the similarity score less than 15%. The project report has been approved as it satisfies the academic requirements in respect of the project work prescribed for the said Degree.

Signature of the Guide	Signature of the Director
Name of the Guide	Name of the Director
Dr. Jay Bharateesh Simha	**Dr. Shinu Abhi**

External Viva Panelists

Names of the Examiners

1. Harsh Vardhan, Chief Digital Technology Architect, Capgemini
2. Pradeepta Mishra, Director – AI, L&T InfoTech
3. Rajib Bhattacharya, Director – Data & Analytics, Cargill

Place: Bengaluru

Date: 20th August 2022

REVA UNIVERSITY
Bengaluru, India

Acknowledgment

I would like to acknowledge the support provided by the founder and Hon'ble Chancellor, **Dr. P Shayma Raju**, Vice-Chancellor, **Dr. M. Dhanamjaya**, and Registrar, **Dr. N Ramesh**.

I would like to express my gratitude to **Dr. Shinu Abhi**, Director of Corporate Training, for providing me significant guidance and assistance throughout the course and my project.

I would like to thank Chief Mentor, **Dr. Jay Bharateesh Simha** for the invaluable assistance he provided as my project guide in helping me to comprehend the concept and execute this project.

I am grateful to Assistant Professor, **Dr. Rashmi Agarwal** and my mentors, **Mr. Akshay Kulkarni** and **Mr. Pradeepta Mishra**, for their great guidance and ideas in learning numerous Text Analytics aspects, as well as for their valuable guidance on a variety of project-related problems.

I would also thank all the mentors and program office at REVA Academy of Corporate Excellence for teaching us awesome methodologies and support throughout the academic year.

REVA UNIVERSITY
Bengaluru, India

Similarity Index Report

This is to certify that this project report titled **"AI-Enabled Automation Solution for Utilization Management in Healthcare Insurance"** was scanned for similarity detection. Process and outcome is given below.

Software Used: Turnitin
Date of Report Generation: 18th August 2022
Similarity Index in %: 8%
Total word count: 6023
Name of the Guide: **Dr. Jay Bharateesh Simha**

Place: Bengaluru
Date: 18th August 2022

Name of the Student: Gaurav Karki
Signature of Student

Verified by:M N Dincy Dechamma

Signature
Dr. Shinu Abhi,
Director, Corporate Training

List of Abbreviations

Sl. No	Abbreviation	Long Form
1	AI	Artificial Intelligence
2	UM	Utilization Management
3	PA	Prior Authorization
4	RPA	Robotic Process Automation
5	NLP	Natural Language Processing
7	AMA	American Medical Association
7	US	United States
8	BOW	Bag-Of-Words
9	TF-IDF	Term Frequency-Inverse Document Frequency
10	MAC	Medicare Administrative Contractor
11	LCD	Local Coverage Determination
12	HCPCS	Healthcare Common Procedure Coding System
13	CMS	Centers for Medicare & Medicaid Services
14	HHS	Health and Human Services
15	EMR	Electronic Medical Records
16	STS	Semantic Textual Similarity
17	RN	Registered Nurses
18	EHR	Electronic Health Record
19	AHIP	America's Health Insurance Plans
20	RTI	Research Triangle Institute
21	DAN	Deep Averaging Network

List of Figure

List of Tables

Abstract

Unstructured forms of text in documents new hurdles as companies move to digitalization by automating more and more processes. The majority of an organization's data are unstructured, and this is a growing trend. Documents are abundant in banking, insurance, health care, and government. Along the whole value chain, enterprises such as healthcare and insurance are making significant progress in implementing business process automation. Business process automation is limited to text extraction to automate document-based processes. Business process automation requires Artificial Intelligence (AI) algorithms that help make decisions, connect information, interpret data, and apply the insights gained to rethink how to make better decisions.

Prior Authorization (PA) is a healthcare procedure that might be improved by using AI. PA is an essential administrative process that is a component of their utilization management systems, and as a condition of coverage, insurers require providers to obtain preapproval for the provision of a service or prescription. This study aims to demonstrate that AI-enabled PA may automate manual processes, hence increasing efficiency, decreasing costs, and freeing up physicians at insurers and Providers should concentrate on complicated situations and patient care and coordination. In response, this may enhance the healthcare experience for both physicians and users in insurance plans.

AI can help combine high-performance computing and deep learning systems for real-time data analysis. Natural Language Processing (NLP) can facilitate insurance claim document processing. This study employs a variety of AI methodologies to evaluate the accuracy and timeliness of the new method's applicability. The accuracy and execution speed of AI models for the PA process are deemed promising.

In this project, an AI-enabled method for validating experimental results is explored and implemented. This study describes the migration of manual procedures to AI-based solutions in order to accelerate them. Numerous potential benefits would result for insurers, providers, and members.

Keywords: Utilization Management, Prior Authorization, Healthcare, Insurance, Claim Processing, Deep Learning, Artificial Intelligence, Automation, Natural language processing

Table of Contents

Chapter 1: Introduction

The expenses of the healthcare system have been spiralling out of control for years, and utilisation management is a crucial method for insurers and providers to guarantee that adequate care is delivered in a cost-effective manner. Utilization management (UM) is the evaluation of medical care based on evidence-based criteria and health payer requirements. In addition to reducing costs, the goal is to give patients with the appropriate treatment, from a shorter duration of stay to enhanced release planning. Once insurers have defined their norms and guidelines, the utilisation management process centres on controlling prior authorization filings via clinical and peer-to-peer evaluations [1].

The influence of UM processes on payer finances, case management, and health plan member and provider satisfaction is direct. Ineffective utilisation control results in annoying care delays and higher operating expenses for the health insurance. However, conventional UM systems place member and provider satisfaction and cost control in opposition. More rules or more comprehensive review processes can reduce the high cost of medical care, but at the expense of member and provider satisfaction. Up until now, there has always been a compromise. Now, insurers have the chance to transform their UM processes through the application of automation and AI. By introducing an AI-based solution for utilisation management, insurers can simultaneously enhance the member and provider experience while reducing operating expenses and medical costs [2].

PA is a primary process of UM. This is conducted before to or at the start of treatment on a particular scenario basis in an effort to eliminate needless services. The selected treatment should be considered provisional and is subject to modification in the future. PA is an examination of a patient's condition and suggested therapy. Its primary objective is to reduce unnecessary, ineffective, or redundant treatments. PA is utilized for regular and urgent referrals, but not for emergency room admissions. The review might take place either before or after admission, but always before treatment begins. In some situations, a physician's directions may not be followed, that could enrage both the medical staff and the patient.

1.1 Utilization Management Process Flow for PA

The UM procedure is intricate. Location, partners, and the medical organization's mission are to determine the requirements. It is hard to draw a process flow that applies universally, but this study can begin by examining the prior authorization review steps as shown in Figure 1.1.

- Verify the patient's eligibility and coverage for the planned treatment.
- Determine how much care the patient needs and whether the planned treatment is medically required by obtaining information about their health.
- If the conditions are satisfied, approve the therapy; if they are not, refuse it.
- The physician may appeal a denial.

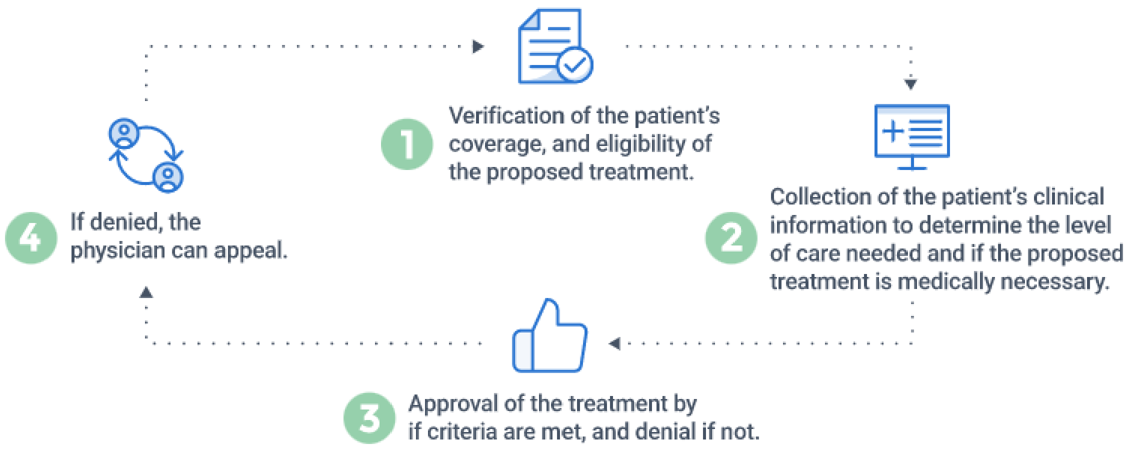

Figure 1. 1 Steps in Utilization Management for Prior Authorisation Process [1]

This study's proposed solution tries to automate the PA procedure. Unlike basic Robotic Process Automation (RPA) systems that just automate individual aspects of the prior authorization process, artificial intelligence may fundamentally alter the way reviewers handle prior authorization requests. If a healthcare provider submits an authorization request, artificial intelligence evaluates this information to the medical necessity criteria to ensure that the patient receives the proper care. If all conditions are satisfied, the request may be automatically authorized, without the requirement for a utilization management reviewer to touch the previous authorization and a clinical evaluation. This can reduce approval times from weeks to hours.

Artificial intelligence may be used to create more efficient PA workflows. Artificial intelligence in tandem with digital and document management systems may aid in organizing data from Electronic Health Records (EHR), emails, policies, clinical treatment, and other sources, drastically reducing low-value, time-consuming human chores like finding, aggregating, and cross-checking information.

NLP, that extracts, analyzes, and integrates unorganized or disordered spoken or textual content, is essential to this attempt. In simpler situations, AI significantly enhances efficient decision, yet in some more complex situations, it may integrate information and provide it to a physician for decision-making.

Chapter 2: Literature Review

PA in medical billing aids the healthcare organization in collecting the correct reimbursement for delivered services, hence lowering denials and subsequent follow-up. According to the results of the Prior Authorization Survey conducted by the American Medical Association (AMA) in December 2021 with the participation of 1,000 practicing physicians, The majority of physicians report an increase in the number of PAs necessary for prescription medications and medical services during the preceding years. In both instances as shown in Figure 2.1, the proportion of physicians reporting this rise ranged from 84% to 84% [3].

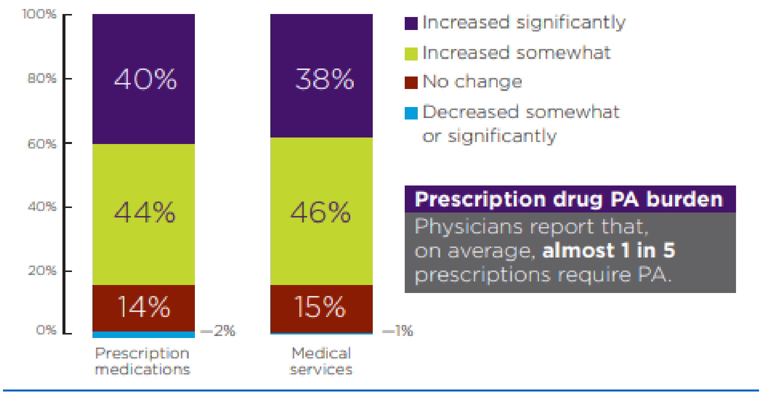

Figure 2. 1 AMA Prior Authorization Physician Survey [3]

According to Susan R. Bailey, M.D., president of the AMA, prior authorization restrictions have been temporarily eased by some commercial health insurers to reduce workload and speed up patients' access to necessary drugs, diagnostics, and treatments as the COVID-19 pandemic begins throughout 2020. The American Medical Association found that most doctors had trouble getting patients access to therapy by the end of 2020, when the U.S. health system was under strain from a weekly record number of new COVID-19 cases [4].

According to a study by America's Health Insurance Plans (AHIP), electronic prior authorizations may enhance care quality by accelerating treatment delivery, lowering provider burden, and boosting the patient experience. The review of over 40,000 transactions showed the impact electronic prior authorization makes in health care, said Denise H. Clayton, research economist of Health Economics and Evaluation at Research Triangle Institute (RTI) International [5].

Researchers analysed trends before and after adoption. They compared trends between providers who used electronic prior permission often and others. After implementing the solutions, 62% of prior authorizations were electronic. Manual prior authorizations almost reduced. Prior authorizations rose 34%. 24% of prior authorization decisions took 48 hours or more before implementation, whereas 17% took 0-2 hours. 15% of prior authorizations took 48 hours or more after implementation, while 33% took zero to two hours, the greatest number pre- or post-implementation. Electronic prior authorization didn't affect approvals [5].

According to McKinsey, AI-enabled PA can automate between 50 and 75 percent of manual processes as shown in Figure 2.2.

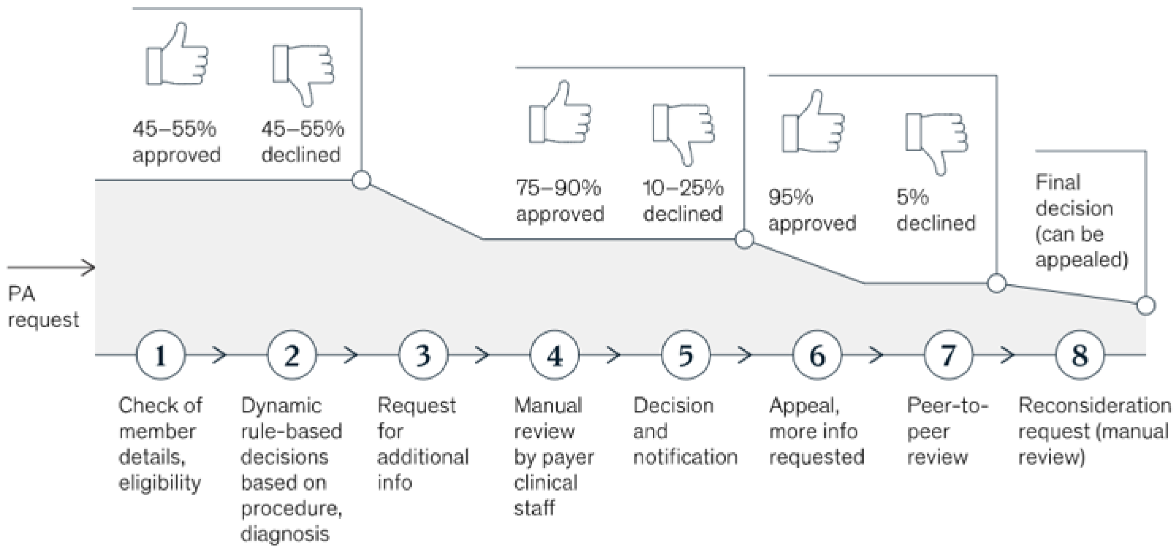

Figure 2.2 The Current PA Workflow [2]

Kumar et al. used Binary Classification on claims data from a major US health insurer to find a considerably greater success rate (accuracy) than earlier methods for identifying claims that needed to be redone [6] . To prevent the submission of false claims, Wojtusiak et al. claimed using attributional methods to foresee potential discrepancies in claims [7]. Claims rework evaluation is distinct from fraud detection, the topic of other research that reported employing ML techniques like rule-based learning and natural language processing to analyze claims [8].

Bag-Of-Words (BOW) [8] and Term Frequency-Inverse Document Frequency (TF-IDF) [9] models have been widely used for text encoding by conventional machine learning algorithms in various text analytics domains, such as the legal sector. Kumar et al. observed that the BOW and TF-IDF model yields better results for recognizing the similarity of legal judgments since just the similarity of the legal phrases, as opposed to all the terms in the dataset, are evaluated [6].

Mandal et al. explored multiple advanced vector representation models for legal document such as Latent Dirichlet Allocation [10] and word embedding along with the TF-IDF model. Word embedding techniques such as Word2vec [11] and Doc2vec [12], that may better capture the semantics of documents, achieved 69% Pearson correlation coefficients in finding similarities among Indian Supreme Court cases, according to the study [13].

Word embedding representation has been used in the legal field for a number of years. However, the results reported in the reviewed existing literature [13], [14], [15] provide a marginal or non-existent improvement over simple BOW models. This study investigates the pre-trained, domain-specific deep learning model utilized for text embedding. After exploring for comparable instances, the newly installed embedding yielded superior results.

Chapter 3: Problem Statement

The PA process is viewed by insurers as an essential element of their utilization management systems. PA is notoriously challenging and time-consuming for both insurers and providers. This study attempts to address the difficulty posed by the typical PA process for insurers.

The challenge is as a result of insurers spending time manually conducting prior authorization verification procedures, the overall cost of care may increase and reimbursement may be delayed. Occasionally, the reason for the delay is that the provider improperly completed the paperwork or omitted vital information. As a result of the PA process, there is a delay in patient care.

This study investigates the feasibility of applying several NLP approaches, notably text similarity algorithms, to develop an optimal model capable of understanding the PA instances semantically and verifying the rules required for PA approval or rejection quickly and accurately. As this would be an automated approach, it would aid insurers in overcoming the obstacles.

Chapter 4: Objectives of the Study

The objective of the study is to develop an AI-enabled solution, consisting of a text similarity model based on NLP and decisions based on semantic analysis, to make utilization management a streamlined and collaborative process. This objective can be divided into sub-objectives:

1. To automate the prior authorization process in order to enhance the frequency of successful auto-decisions and prioritize clinical reviews appropriately.

2. To expedite patient care while increasing member satisfaction.

3. Increase insurers' control over patients' medical claims while simultaneously achieving administrative efficiency and enhancing the provider experience by unlocking unstructured data to contextualize authorization requests.

Chapter 5: Project Methodology

The method proposes in this work offers an approach for screening PA cases based on local coverage determinations, that are decisions made by a Medicare Administrative Contractor (MAC) whether to approve a case or not. This is an NLP-based model to help insurers by screening the PA cases that give the most similar information and ranking them based on the similarity score in comparison with the Local Coverage Determinations (LCDs) criterions. The Proposed solution consists of the following components as shown in Figure 5.1.

1. PA case
2. Requested process id and Healthcare Common Procedure Coding System (HCPCS) code
3. LCDs Criterions
4. Lexicons
5. Text Similarity Check

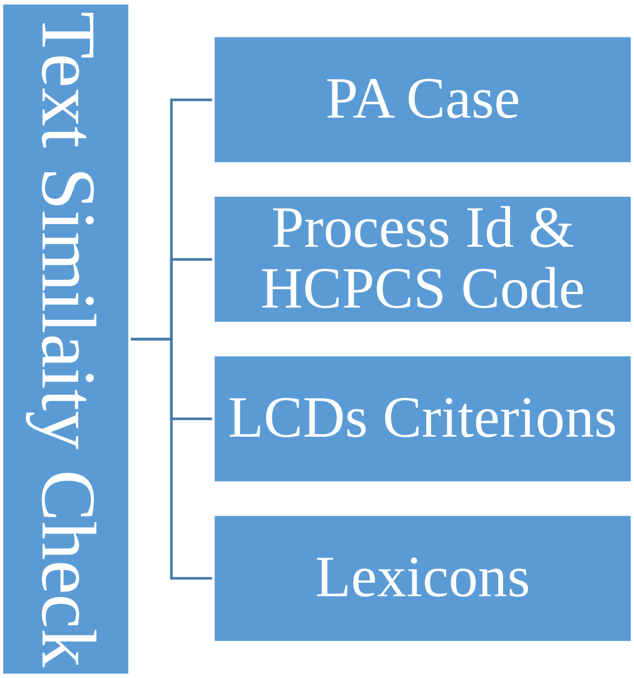

Figure 5. 1 The Proposed Solution's Components

PA cases are collected through any channel between Insurer and provider, including Electronic Medical Record (EMR). Then, getting prepared three JSON files containing all the necessary information to be utilised as input for the text similarity model.

1. The first file contains the rules from the LCD criteria used to approve the procedure in relation to the HCPCS code.

2. The second file contains the lexicons associated with each rule; these lexicons aid in the identification of required information from PA cases.

3. The third file contains both the requested process id and the HCPCS code needed to verify the eligibility of PA case.

The model retrieves text from the lexicon file and PA case and then executes all of the steps outlined in the following Figure 5.2.

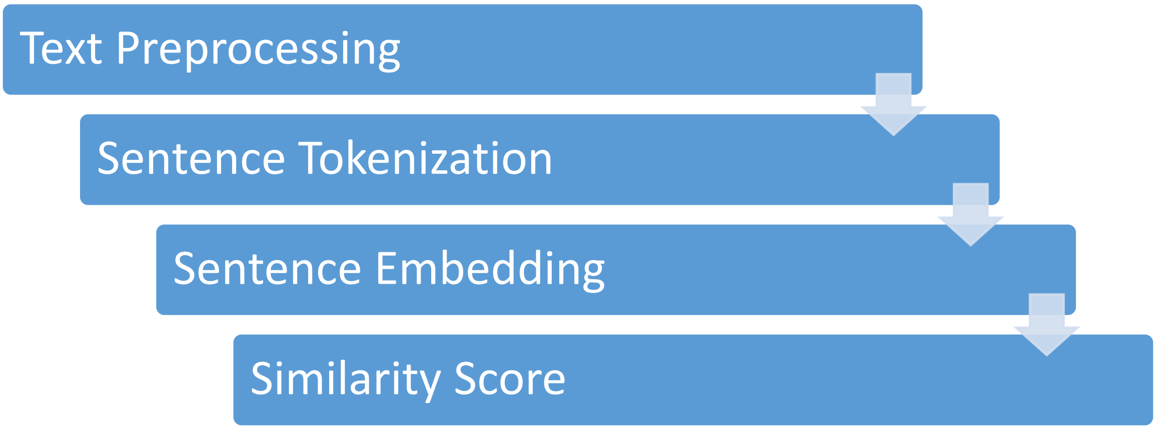

Figure 5. 2 Text Analysis Workflow

5.1 Challenges

There are a number of challenges in this study, that are outlined below:

1. Getting a PA case is one of the main challenges to achieving the goal of the study. A payer can get a PA case through EMR, but in this study dummy PA cases are used.

2. To set the rules for the lexicon Consequently, this solution must go through as many PA cases as possible to comprehend the rules indicated in PA cases.

Chapter 6: Resource Requirement Specification

There are two distinct sorts of required resources.

1. Data Resources

2. Technical Resources

6.1 Data Resources

In this study, particular requirements for open source data are considered; these requirements are listed below.

6.1.1 Process Selection

There are many processes under PA and in this research, we narrowed the scope of the project to include only one process that is powered mobility devices.

6.1.2 Guidelines

The study pertains to powered mobility devices, therefore local coverage determinations, that are decisions made by a Medicare Administrative Contractor (MAC)., are referenced as guidelines. And these rules are collected from the medical coverage information shared by the Centers for Medicare & Medicaid Services (CMS), a division of the United States Department of Health and Human Services (HHS) in United States of America. The Process Id for powered mobility devices as mentioned at CMS is L33789.

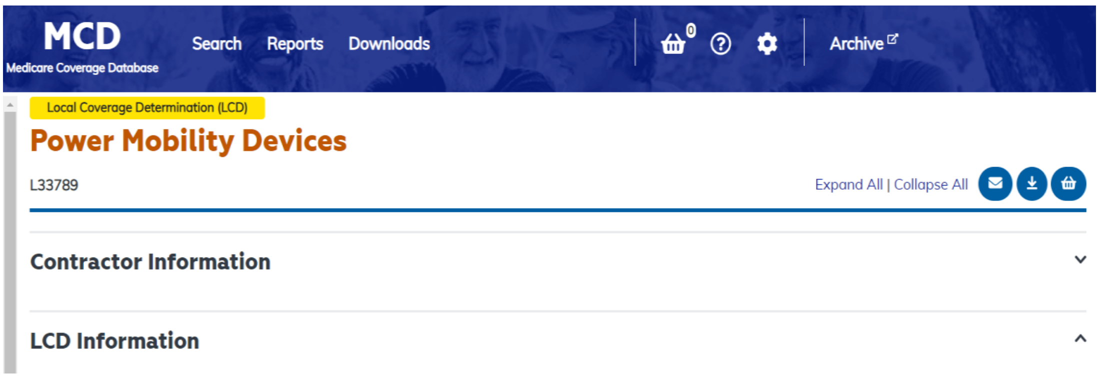

Figure 6. 1 Local Coverage Determinations for Power Mobility Devices

All HCPCS codes can be extracted from the same database. Few HCPCS codes for powered mobility devices are mentioned below:
K0801, K0820, K0816, K0815, K0814, K0813, K0812, K0808, K0807, K0806, K0802

6.1.3 PA Cases

This is an essential necessity for data resources, as this research is conducted on PA cases. In light of this, dummy instances pertaining to the process for power mobility devices have been produced for this study.

6.2 Technical Resources

The following are the requirements for this resource.

1. Python environment with required libraries
    - NLTK
    - Pandas
    - Numpy
    - Json
    - Tensorflow
    - Scipy
    - re
2. Semantic Textual Similarity (STS) benchmark data for evaluation
3. Pre-trained model - Universal Sentence Encoder

Chapter 7: Software Design

This study intends to automate the PA process; thus this study must first comprehend the manual procedure.

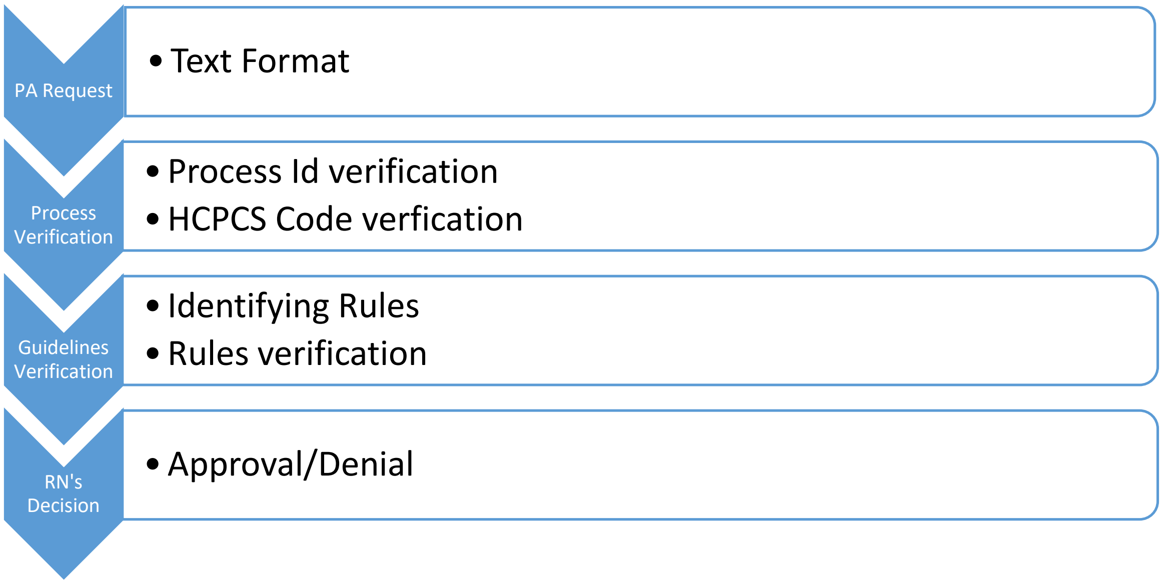

Figure 7. 1 Manual Workflow for PA Process

This study's manual procedure, as shown in the Figure 7.1, aids in the development of an AI-enabled solution.

7.1 Proposed AI-Enabled PA Process

PA's manual workflow is being streamlined, that is, making it a process that can be done again and again. Therefore, AI can contribute to the acceleration of this process.

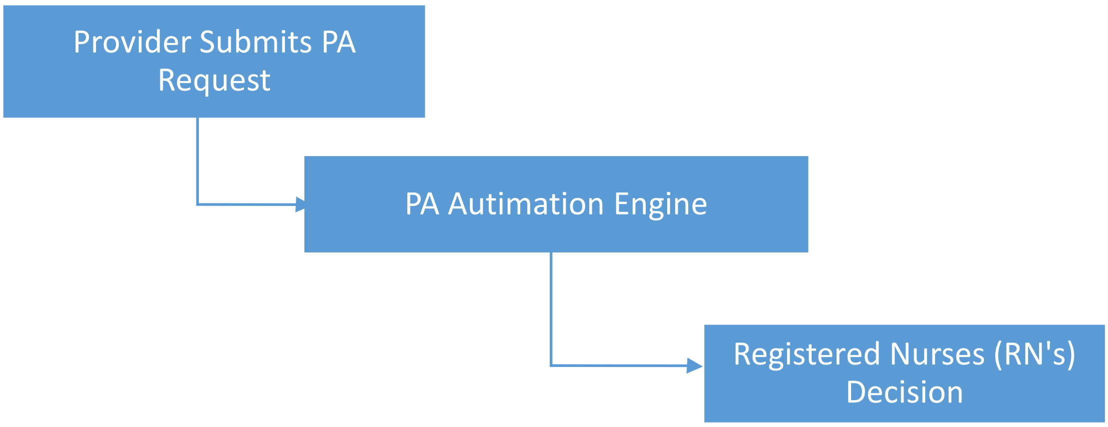

Figure 7. 2 Proposed AI-Enabled PA Process

As depicted in Figure 7.2, the suggested AI-enabled solution consists primarily of three phases. First, the provider submits a PA request, then the PA automation engine examines the case and give the necessary information from the text to assist Registered Nurses (RNs) in making an approval or denial determination. Now examine the automation engine's underlying framework.

7.2 Automation Engine

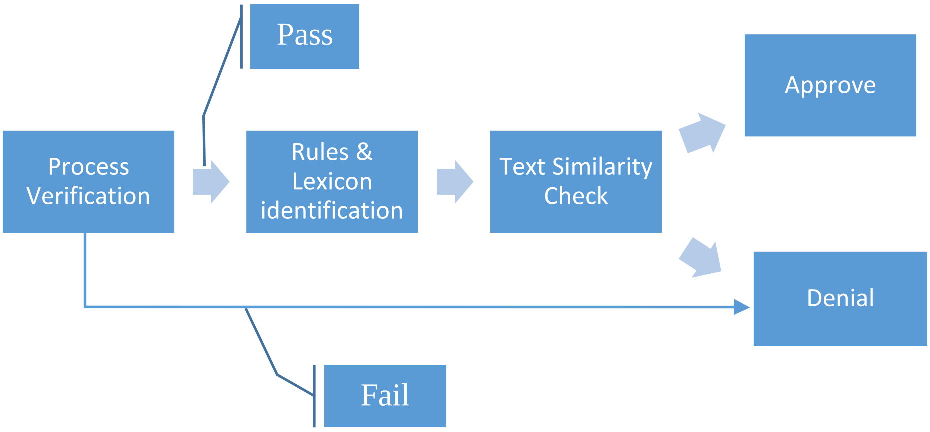

Figure 7. 3 Automation Engine Process

There are three stages involved in the operation of an automation engine as shown in Figure 7.3. The first stage involves process verification; the second stage identifies rules associated with process id and HCPCS code, as a result, finalises the lexicons for each rule. and the third stage examines textual similarities. If the process verification phase is successful, the subsequent phases are enabled; if it is failed, a case of rejection is possible. Each step is elaborated addressed separately.

7.2.1 Process Verification

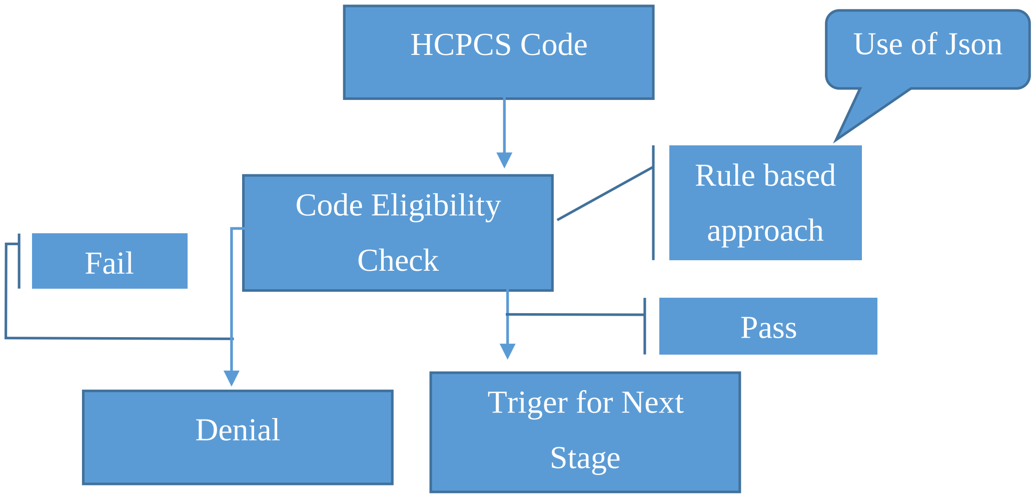

Figure 7. 4 Process Verification Process

The Figure 7.4 illustrates that the process verification procedure begins with the input of the HCPCS code. Therefore, the automation engine employs a rule-based method to cross-check with json files containing the process id and corresponding codes. If this eligibility check is successful, the procedure advances to the subsequent level; otherwise, the case is denied.

7.2.2 Rules and Lexicon Identification

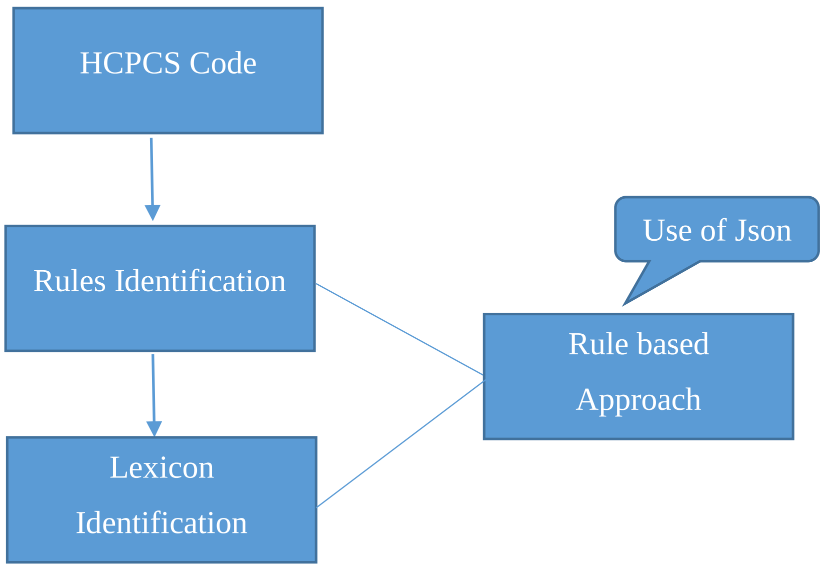

Figure 7. 5 Rules and Lexicon Identification Process

The Figure 7.5 demonstrates that HCPCS code facilitates the identification of rules using a rule-based methodology. Consequently, lexicon is identified relative to each rule using the same method.

7.2.3 Text similarity check

This phase is the central concept of AI utilisation in this study. Textual analysis comes into play here.; it aids with text comprehension. As shown in Figure 7.6, Text is provided as input at this stage, after that text preparation occurs. Then, the sentence is tokenized so that the model can convert it into a vector, that is the sentence embedding process. The identical procedure of sentence embedding was used to the lexicon identified in the previous step. Now, after obtaining two sets of vectors, one from lexicons and the other from text, cosine similarity is used to generate similarity metrics. After receiving ranking evidence, RNs can make a decision based on this evidence.

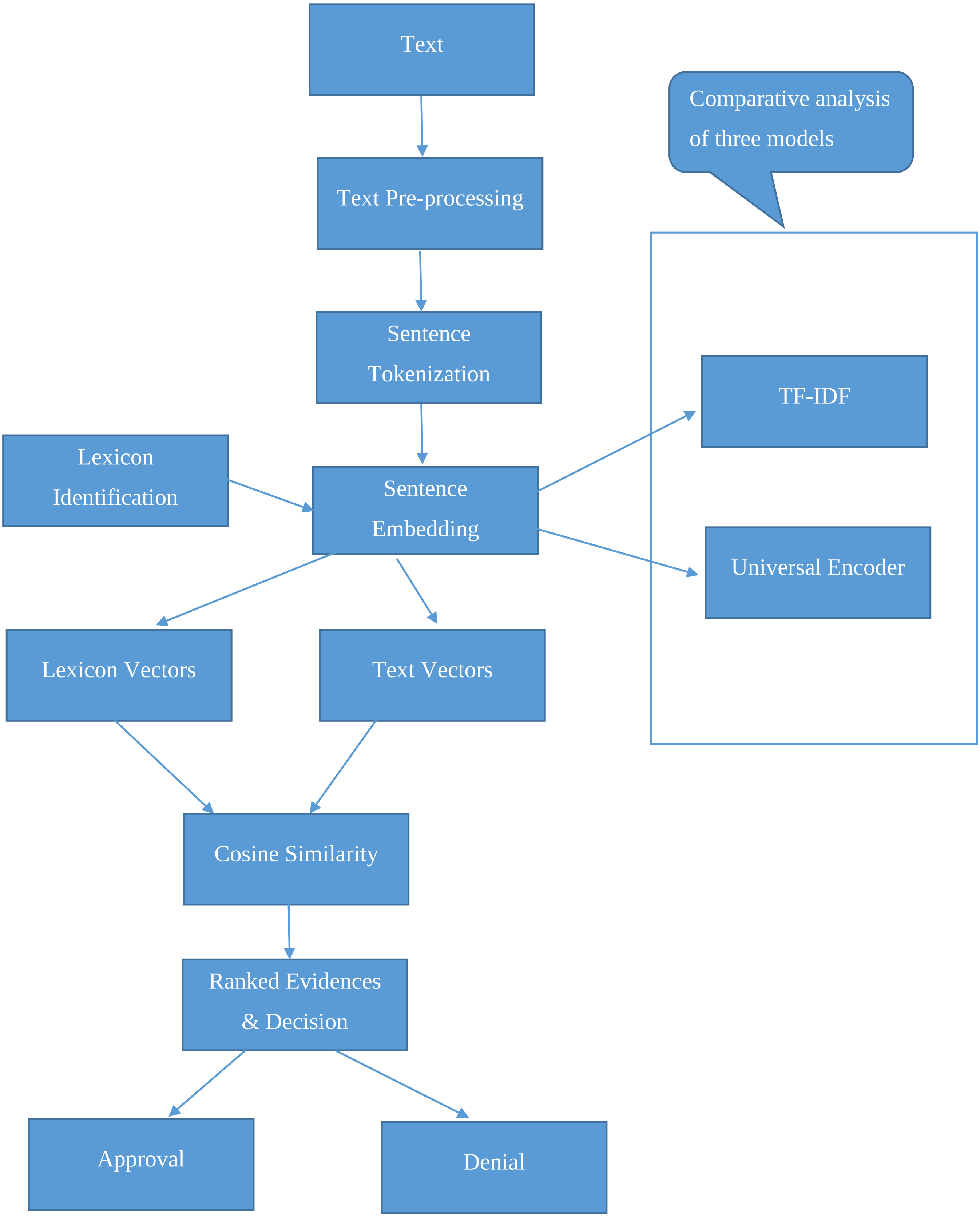

Figure 7. 6 Text Similarity Check

Chapter 8: Implementation

8.1 Data Preparation

In an effort to create an AI-enabled solution. The first and possibly most crucial phase is data preparation, because the quality of the solution depends on the quality of the data used. On the payer side, data is collected in accordance with rules. All essential information is kept in JSON format in a configuration folder containing data collected from insurers in accordance with their requirements.

8.1.1 Configuration Folder

1. **Process File –** This file includes the process id and all HCPCS codes as shown in Figure 8.1.

```
"L33789": {
    "process_desc": "PWC",
    "scope_hcpcs_codes": [
        "K0000", "K0013", "K0814", "K0815", "K0816", "K0820", "K0821", "K0822", "K0823", "K0824",
        "K0825", "K0826", "K0827", "K0828", "K0829", "K0835", "K0836", "K0837", "K0838", "K0839",
        "K0840", "K0841", "K0842", "K0843", "K0848", "K0849", "K0850", "K0851", "K0852", "K0853",
        "K0854", "K0855", "K0856", "K0857", "K0858", "K0859", "K0860", "K0861", "K0862", "K0863",
        "K0864", "K0868", "K0869", "K0870", "K0871", "K0877", "K0878", "K0879", "K0880", "K0884",
        "K0885", "K0886", "K0890", "K0891", "K0898"
    ]
}
```

Figure 8. 1 Process File Containing All HCPCS Code

2. **Rules File –** This file contains descriptions of all rules linked with the specified process ID as shown in Figure 8.2.

```
"L33789": {
    "A": "<p>The beneficiary has a mobility limitation that significantly impairs his/her   ability
    to participate in one or more mobility-related activities of daily living (MRADLs) such as
    toileting, feeding, dressing, grooming, and bathing in customary locations in the home.</p><p><b
    class=\"text-primary\"> NOTE: All of the basic coverage criteria (A-C) should met</b></p>",
    "B": "<p>The beneficiary's mobility limitation cannot be sufficiently and safely resolved by the
    use of an appropriately fitted cane or walker.</p><p><b class=\"text-primary\"> NOTE: All of the
    basic coverage criteria (A-C) should be met</b></p>",
```

Figure 8. 2 Rules File Containing All Rules Descriptions

3. **Lexicon File -** This file is comprised of all lexicons related with each rule. Similarly, more lexicons can be introduced as shown in Figure 8.3.

```
"L33789": {
    "A": [
        "The beneficiary has a mobility limitation that significantly impairs his/her ability to
        participate in one or more mobility -related activities of daily living (MRADLs) such as
        toileting, feeding, dressing, grooming, and bathing in customary locations in the home",
        "A mobility limitation prevents the beneficiary from accomplishing an mobility -related
        activities of daily living (MRADLs) entirely",
        "A mobility limitation places the beneficiary at reasonably determined heightened risk of
        morbidity or mortality secondary to the attempts to perform an mobility - related activities
        of daily living (MRADLs)",
        "A mobility limitation prevents the beneficiary from completing an mobility -related
        activities of daily living (MRADLs) within a reasonable time frame",
        "can not / unable to perform mobility -related activities of daily living (MRADLs)",
        "significantly impairs his/her ability to participate in one or more mobility -related
        activities of daily living such as toileting, feeding, dressing, grooming, and bathing",
        "home / self management heightened risk of morbidity",
        "risk of mortality in an attempt to perform mobility -related activities of daily living
        (MRADLs)"
    ],
```

Figure 8. 3 Lexicon File Containing All Lexicons

4. **Flow File:** This file contains the flow structure for all rules being implemented for all circumstances. Then the process id "L33789" gets assigned and within that dictionary, there are two keys named "start rule" and "rules" with their corresponding values: rule "A" and a dictionary containing all rules from "A" to "O." Each rule compares the input code to the provided list of codes. Therefore, the rule's application is verified by the presence of input code in the list as shown in Figure 8.4.

```
"L33789": {

    "start_rule": "A",

    "rules": {
        "A": {
            "rule_type": "indication",
            "on_pass": "B",
            "on_fail": "end",
            "flag_fail": "denial"
        },

        "B": {
            "rule_type": "indication",
            "on_pass": "C",
            "on_fail": "end",
            "flag_fail": "denial"
        },
```

Figure 8. 4 Flow File

8.3 Implementation of The Proposed Solution

The AI-enabled solution for PA requires two components for implementation. The provider provides PA cases in text format with HCPCS codes. Once a PA case containing the requested HCPCS code is received, the automation engine is engaged. According to the design, the automation engine consists of three stages.

1. **Process Verification:** During this phase, the HCPCS code is utilised to validate the process id using process data from the process file.

```python
#to get process id based on HCPCS code
def get_process_id(client_hcpcs):

    for pid ,pval in process_data.items():
        if client_hcpcs in pval['scope_hcpcs_codes']:
            return pid

    return None
```

Figure 8. 5 Process Verification Function

2. **Rules & Lexicon Identification:** At this stage, all rules are identified, and then the lexicon associated with each rule is accessed.

```python
rules_in_scope = []

process_codes = process[curr_process_id]['scope_hcpcs_codes']
process_flow = flow[curr_process_id]['rules']

for k, _ in process_flow.items():

    # determine the codes (cpt + hcpcs) that a given rule is applicable on #

    # if no codes are mentioned in json config file, assume that the rule is valid on all process codes
    if process_flow[k].get("codes_cpt") is None and process_flow[k].get("codes_hcpcs") is None:
        codes_in_scope = process_codes

    # merge cpt and hcpcs into a single list
    else:
        codes_in_scope = process_flow[k].get("codes_cpt", []) + process_flow[k].get("codes_hcpcs", [])

    for code in doc_codes:
        if code in codes_in_scope:
            rules_in_scope.append(k)

rules_in_scope
```

```
['A', 'B', 'C', 'D', 'E', 'F', 'J', 'K', 'M', 'N', 'O', 'SWO']
```

Figure 8. 6 Rules Associated to Process ID

3. **Text Similarity Check:** This is the final level of Ai's implementation. As described in the solution design section, a PA case must be formatted as illustrated in Figure 8.7.

```python
f = open ('s1.txt', 'r', encoding='utf-8')
og_string = f.read()
og_string
```

Python

```
"Patient is being evaluated for a motorized wheelchair. He presents today with Diabetes Mellitus Typa li, HTN, Osteoarthrits, and Gait Instability, Hie symptoms
include Joint pain, fatigue, upper extremity weakness. He has a history of falls, once resulting in left ankle fusion. He's limited In his ability to participate
in all mobility related activities of daily living in the home setting. He has difficulty accessing the kitchen for meal preparation, and accessing the bathroom
In a timely manner for tolleting/ hygiene needa. He la unable to safely or effectively use a cane or walker for the distance needed in the home due to fatigue,
joint pain, and numbness in RLE. He is unable to self-propel an optimally configures manual wheelchair due to upper extremity weakness and arthritic hand pain. A
motorized scooter would not meet the patient's needs in the home due to lack of | maneuvering apace between rooms. The patient's home does provide adequate
access between rooms for a motorized wheelchair, The power chair is recommend to reduce fall risk and to increase Independence in completing mobility related ;
ADL's safely in a more timely manner. He has the physical and mente! capabilities to safely operate the chair in the home and he will use it on a regular basis.
The chair will require the following accessories: Batteries to power the wheel chair. Adjustable height arm rest are recommended for upper extremity positioning
and support. A swing away joystick mountis recommended – to move joystick cut of way for safe transfers. ROS Patient reports no fever and no significant weight
gain. He reports no vision change. He reports no difficulty hearing. He reports no nose/sinus problems. He reports no ora! abnormalities. He reports no chest
pain. He reports no cough. Fle reporta _ NO abdominal pain. He reports no difficulty urinating. He reports no muscle weakness, no arthralgias/joint pain, and no
swelling : In the extremities. Ha reports no abnormal mole and no rashes. He raports no dizziness. He reports no depression. He reports no fatigue. He reports no
swolten glands. He reports no runny nose.  "
```

Figure 8. 7 Text Format

The following are the steps that need to be taken in order to finish this stage.

STEP 1: Create a data frame containing tokenized sentences from the original text string.

STEP 2: Text Pre-processing - to prepare the text data for model development. It is the initial phase of NLP initiatives. Some pre-processing steps include:

- Removing punctuations like . , ! $( ) * % @
- Keeping only alphabetical characters
- Lower casing
- Striping unnecessary whitespaces
- Tokenization
- Lemmatization

```python
import re
import nltk
from nltk.stem import WordNetLemmatizer


def clean_sentence(s, lemmatize=False):
    """
    lowercase, keep only alphabetical characters
    optional: lemmatize the tokens in sentence
    """

    # lowercase
    s = s.lower()

    # keep only alphabetical chars
    s = re.sub(r"[^a-z ]", " ", s)

    # strip unnecessary whitespace
    s = re.sub(r" +", " ", s).strip()

    if lemmatize:
        lemma = WordNetLemmatizer()
        s = " ".join([lemma.lemmatize(token) for token in nltk.word_tokenize(s)])

    # sanity check
    s = " ".join([token for token in nltk.word_tokenize(s)
                  if len(token) > 1 and len(token) < 15])

    return s
```

Figure 8. 8 Text Pre-Processing Function

Here Several processes, such as text preparation, are utilised in this study, as depicted in the Figure 8.8. The original sentences were implemented in the data frame.

STEP 3: Once text pre-processing is complete, sentence embedding is performed to generate a dictionary of pre-processed sentences and their corresponding vectors. Vectors are derived using a model. This model is founded on NLP principles. In this study, two NLP strategies were utilised to determine the most effective method for achieving our objective.

8.4 TF-IDF

In this approach, the frequency of words is rescaled according to their overall frequency in all texts, which penalizes common, ubiquitous words like "the" that appear often throughout all texts. TF-IDF measures how crucial a specific term is to the overall meaning of a text. Multiplying two separate metrics yields a document word's TF-IDF.

The Term Frequency (TF) of a document's words. There are numerous methods for calculating this frequency, the simplest of that is a simple count of the occurrences of a word in a document. Then, there are further methods for adjusting the frequency. For instance, as Equation 8.1 describes, by dividing the raw count of occurrences of a word by the document's length or by the raw frequency of the document's most frequent word.

$$TF(i,j)= n(i,j) \,/\, \Sigma\, n(i,j) \qquad (8.1)$$

Where,

$n(i,j)$ = number of times nth word occurred in a document

$\Sigma\, n(i,j)$ = total number of words in a document.

The inverse document frequency (IDF) of a given word across a collection of documents. This reflects the frequency of a word in the entire document set. The closer a term is to 0, the more frequent it is. This metric can be determined by dividing the total number of documents by the number of documents containing a specific word. This metric can then be calculated using the logarithm.

Therefore, this number approaches 0 if the term is prevalent and appears in several documents. Alternatively, it approaches 1 Multiplying these two numbers yields the TF-IDF score of each word in a document. The higher the score, the more pertinent the word is to the document. In mathematical terms, the TF-IDF score is calculated according to Equation 8.2.

$$IDF=1+log(N/dN) \qquad (8.2)$$

Where,

N =Total number of documents in the dataset

dN =total number of documents in that nth word occur

8.5 Universal Sentence Encoder

A significant amount of work is expended in machine learning research to convert data into vectors. Word2vec and Glove [16] accomplish this by turning a word into a vector. Therefore, the vector corresponding to "cat" will be closer to "dog" than to "eagle." While embedding a sentence with its words, however, the complete sentence's context must be captured in that vector. The "Universal Sentence Encoder" comes into play at this point.

The embedding generated by the Universal Sentence Encoder [17] model especially transfer learning to the NLP tasks. It is trained on a number of data sources in order to acquire skills for a vast array of tasks. The sources include Wiki, web media, online question-and-answer pages, and forums. The input is variable-length English text, while the outcome is a 512-dimensional vector.

Typically, sentence embedding was derived by averaging the embedding of all the words in the phrase; however, this method had limitations and was unsuitable for detecting the true semantic meaning of a sentence. The Universal Sentence Encoder makes sentence-level embedding effortless. It is available in two variants, one trained with the Transformer encoder and the other with the Deep Averaging Network (DAN). In terms of computer resource requirements and accuracy, there is a trade-off between the two. While the one with the Transformer encoder is more precise, it requires more computation. The variant with DAN encoding is computationally less expensive and slightly less precise. This study utilizes the transformer encoder variant.

Chapter 9: Testing and validation

Text similarity is an unsupervised task, and there is no clear way to objectively analyze the results. Evaluation is dependent upon the intended use. Calculating the similarity score concludes the testing and validation process. Similarity measure can be used to calculate the distance between the similarity of two objects being compared and improve the accuracy of information retrieval. A common method for matching comparable documents is based on the highest number of shared terms across the documents. However, this strategy has an inherent drawback.

In other words, as the size of a document increases, the frequency of similar words tends to rise, even if the documents cover different subjects. This basic problem in the "count the frequent words" or Euclidean distance technique is overcome by the cosine similarity metric.

9.1 Cosine Similarity

Cosine Similarity is a measure of the similarity between two vectors derived from the cosine angle product of the vectors being compared [18].

Figure 9.1 depicts the concept of assessing document similarity using the cosine angle, where vector coordinates indicate the compared documents and the cosine degree between vectors represents the similarity degree. If, according to the cosine principle, cosine 0° equals 1 and less than 1 to the value of another angle, then two vectors are similar if their cosine similarity value is 1 [19].

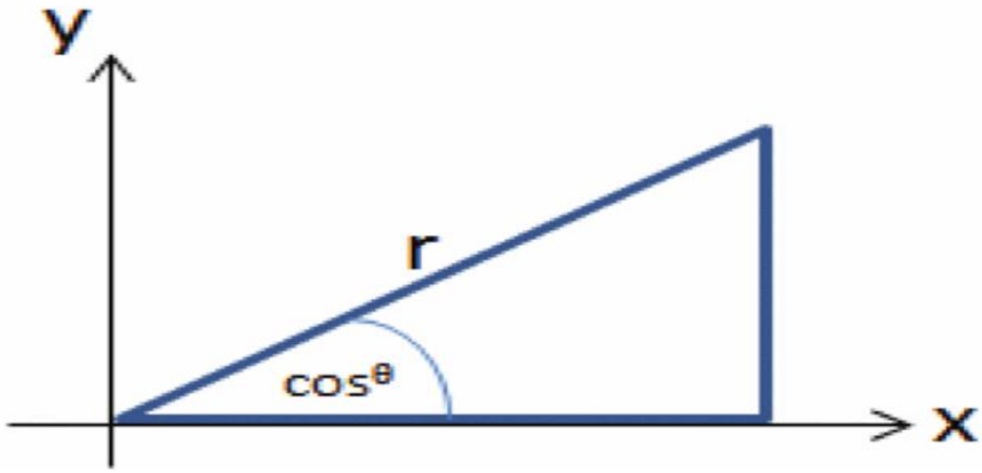

Figure 9. 1 Cosine Similarity [19]

Calculation of the cosine similarity coefficient using the formula as shown in Equation 9.1.

$$\boldsymbol{cos} \propto = \frac{\boldsymbol{A} \times \boldsymbol{B}}{|\boldsymbol{A}| \times |\boldsymbol{B}|} = \frac{\sum_{i=1}^{n} \boldsymbol{A}_i \times \boldsymbol{B}_i}{\sqrt{\sum_{i=1}^{n} (\boldsymbol{A}_i)^2} \times \sqrt{\sum_{i=1}^{n} (\boldsymbol{B}_i)^2}} \tag{9.1}$$

As discussed, in this study two NLP techniques has been used. And testing and validation purpose only few rules have been considered that are Rule “A”, Rule “B” and Rule “C”.

Rule “A” says one or more mobility-related activities of daily living, are severely hindered because of the beneficiary's mobility limitation.

Rule “B” says a properly adjusted cane or walker would not help the beneficiary with his/her mobility issues to an acceptable degree.

Rule “C” says there is insufficient upper extremity function for the beneficiary to propel a properly equipped manual wheelchair indoors.

Results that are from the TF-IDF technique are mentioned in the Table 9.1.

Rule name	**Total no. of Matches**	**Top matching sentence from PA text**	**Highest score**
A	5	he limited in his ability to participate in all mobility related activities of daily living in the home setting	0.491431
B	6	he la unable to safely or effectively use cane or walker for the distance needed in the home due to fatigue joint pain and numbness	0.2933
C	9	he is unable to self-propel an optimally configures manual wheelchair due to upper extremity weakness and arthritic hand pain	0.3841

Table 9. 1 Results from the TF-IDF Technique

Results that are from the Universal Sentence Encoder Technique are mentioned in the Table 9.2.

Rule name	**Total no. of Matches**	**Top matching sentence from PA text**	**Highest score**
A	7	he limited in his ability to participate in all mobility related activities of daily living in the home setting	0.72
B	5	he la unable to safely or effectively use cane or walker for the distance needed in the home due to fatigue joint pain and numbness in rle	0.70
C	8	he is unable to self-propel an optimally configures manual wheelchair due to upper extremity weakness and arthritic hand pain	0.71

Table 9. 2 Results from the Universal Sentence Encoder Technique

Chapter 10: Analysis and Results

In this study, analysis is done by STS Benchmark. This Benchmark offers an empirical assessment of the degree to which similarity ratings obtained by sentence embedding correspond to human judgments. The benchmark necessitates that systems produce similarity scores for an assortment of sentence pairs. The Pearson correlation coefficient is then applied to compare the quality of algorithm similarity scores to human judgments. The statistical method of the Pearson correlation coefficient [20] is commonly used in economics for purposes like trend analysis and classification. Other potential domains of use have been discussed in recent years' literature. Using it, one can determine if strongly two variables are related to one another along a linear axis. And p-value is calculated to check statistically significance. If the correlation coefficient were indeed zero, then the current result would have been seen with a probability equal to the P-value (null hypothesis). A correlation coefficient is considered statistically significant if its associated probability is less than 5%.

10.1 TF-IDF

Pearson correlation coefficient = 0.2340

p-value = $1.015e^{-19}$

10.2 Universal Sentence Encoder

Pearson correlation coefficient = 0.83

p-value = 0.0

So, the result from the both technique says Universal Sentence Encoder has high correlation coefficient and p-value compare to TF_IDF.

10.3 Similarity Visualization

Similarity is visualised using a heat map. The graph is a 31x52 matrix and the color of each element [i, j] is determined by the dot product of the embedding for sentence i and j.. Figure 10.1 shows heat map for TF-IDF and Figure 10.2 shows heat map for Universal Sentence Encoder.

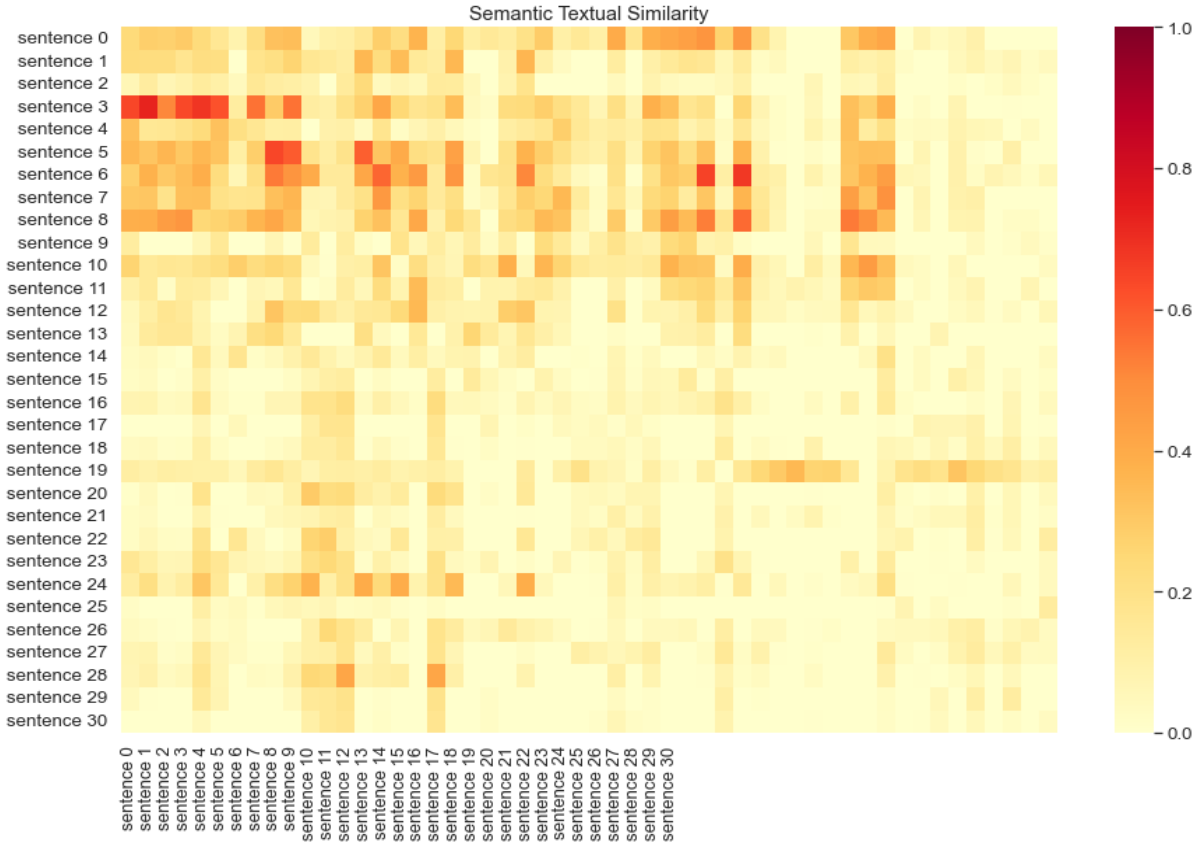

Figure 10. 1 Similarity Visualization for TF-IDF

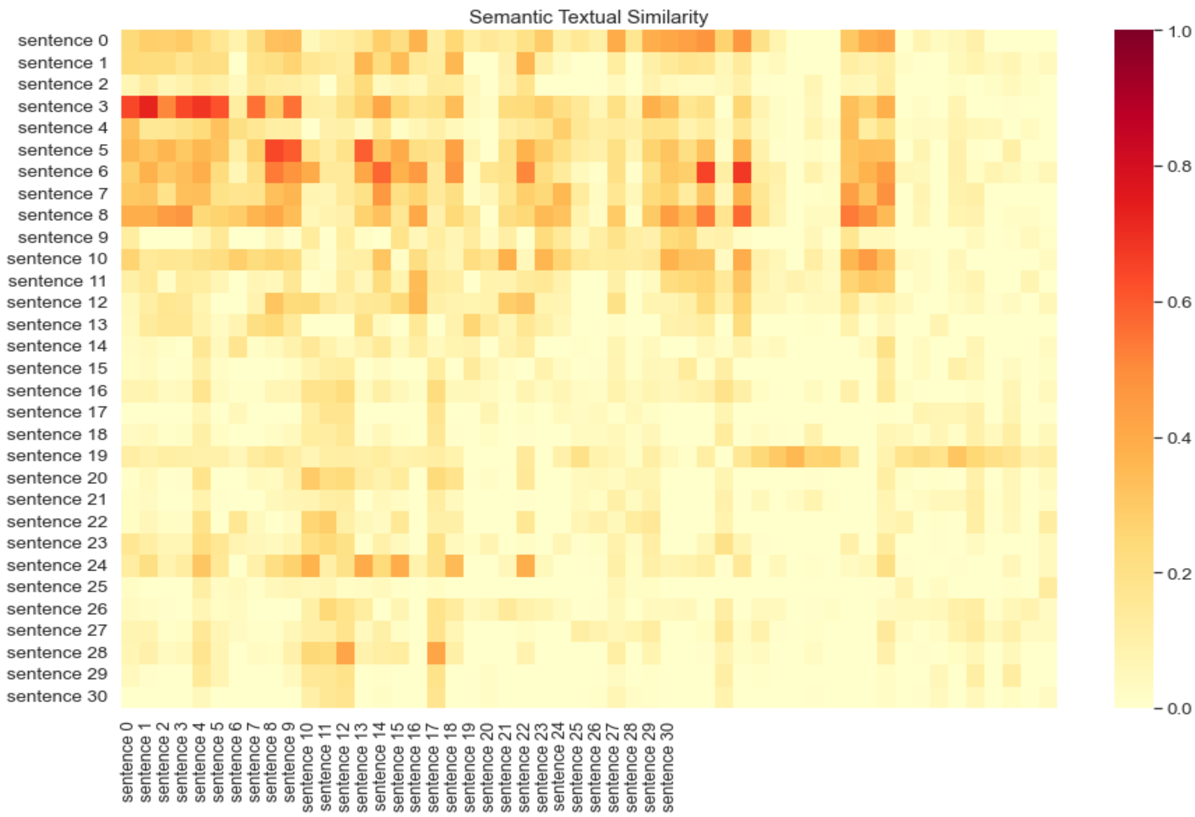

Figure 10. 2 Similarity visualization for Universal Sentence Encoder

Chapter 11: Conclusions and Future Scope

The facts presented in the previous section make it plainly clear that Universal Sentence Encoder's tactics are superior to those of TF IDF. Using the Pearson correlation coefficient, the assessment standard allows for differentiation between the various methodologies. The similarity visualization generates outputs with varying degrees of color intensity that are quite similar. This conclusion is based on the information gathered for the purpose of this study. On the other hand, the evaluation benchmark makes it obvious that Universal Sentence Encoder approaches can still tackle the challenge even if the difficulty of the text increases. Even though this study is conducted on PA cases, that are part of the healthcare industry, complexity management is always a concern. This methodology aids in reaching the objectives of the study.

Access to necessary care for patients is frequently delayed as a result of prior authorizations, that may drive patients to abandon their treatment due to the waiting period or other complications related with prior authorization. This work provides a viable method for resolving the issue, as it proposes a method for streamlining AI that minimizes treatment delays and disruptions by reducing the requirement for prior approval.

This solution, that is an integral part of the End-to-End Prior Authorization process, eliminates human work that is time-consuming and prone to error. Therefore, the pre authorization team can maximize the health system's capacity to provide faster and better care. Patients, healthcare providers, and insurers, as well as any other parties engaged in the process, can all benefit from an efficient utilization management programme. These are the adverbial complements for each:

1 Patients gain from decreased treatment costs, more treatment efficacy, and fewer refused claims.

2 Fewer denied claims, reduced costs, more effective treatments, improved data, and more efficient resource utilization are all beneficial to the health care industry.

8.1 Future Scope

The results of this study provide businesses with a list of recommendations about prior authorization and the selection of a health plan that may help enhance patient outcomes during the course of their treatment.

Insurers should inquire about the prior authorization requirements of such programme and the potential effects these rules may have on their workforce. The second step for insurers is to ask employees about their prior authorization experiences. This study motivates us to solve this problem by approaching it as a classification problem.

This study serves as proof-of-concept for the implementation of NLP approaches as an alternative to the manual PA procedure. The proposed data-driven method yielded encouraging results and contributes to the formation of a foundation for tackling end-to-end PA process utilization management.

Bibliography

[1] T. M. Wickizer and D. Lessler, “Utilization management: Issues, effects, and future prospects,” *Annu. Rev. Public Health*, vol. 23, pp. 233–254, 2002, doi: 10.1146/ANNUREV.PUBLHEALTH.23.100901.140529.

[2] “AI ushers in next-gen prior authorization in healthcare | McKinsey | McKinsey.” https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/ai-ushers-in-next-gen-prior-authorization-in-healthcare (accessed Aug. 10, 2022).

[3] A. Medical Association, “Prior Authorization Physician Survey Update | AMA,” 2022, Accessed: Aug. 10, 2022. [Online]. Available: https://www.ama-assn.org/system/files/prior-.

[4] “Most physicians had little relief from prior authorization as COVID cases soared | American Medical Association.” https://www.ama-assn.org/press-center/press-releases/most-physicians-had-little-relief-prior-authorization-covid-cases (accessed Aug. 10, 2022).

[5] “Electronic Prior Authorizations Impact Quality, Provider Burden.” https://healthpayerintelligence.com/news/electronic-prior-authorizations-impact-quality-provider-burden (accessed Aug. 10, 2022).

[6] M. Kumar, R. Ghani, and Z. S. Mei, “Data mining to predict and prevent errors in health insurance claims processing,” *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 65–73, 2010, doi: 10.1145/1835804.1835816.

[7] J. Wojtusiak, C. Ngufor, J. Shiver, and R. Ewald, “Rule-based prediction of medical claims’ payments: A method and initial application to medicaid data,” *Proc. - 10th Int. Conf. Mach. Learn. Appl. ICMLA 2011*, vol. 2, pp. 162–167, 2011, doi: 10.1109/ICMLA.2011.126.

[8] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, “The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature,” *Decis. Support Syst.*, vol. 50, no. 3, pp. 559–569, 2011, doi: 10.1016/J.DSS.2010.08.006.

[9] J. Lam, Y. Chen, F. Zulkernine, and S. Dahan, “Detection of Similar Legal Cases on Personal Injury,” *IEEE Int. Conf. Data Min. Work. ICDMW*, vol. 2021-December, pp. 639–646, 2021, doi: 10.1109/ICDMW53433.2021.00084.

[10] D. Blei, A. Ng, M. J.-J. of machine L. research, and undefined 2003, “Latent dirichlet allocation,” *jmlr.org*, vol. 3, pp. 993–1022, 2003, Accessed: Aug. 10, 2022. [Online].

Available:
https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf?ref=https://githubhelp.com.

[11] T. Mikolov, K. Chen, G. Corrado, J. D. preprint arXiv:1301.3781, and undefined 2013, "Efficient estimation of word representations in vector space," *arxiv.org*, Accessed: Aug. 10, 2022. [Online]. Available: https://arxiv.org/abs/1301.3781.

[12] Q. Le, T. M.-I. conference on machine, and undefined 2014, "Distributed representations of sentences and documents," *proceedings.mlr.press*, Accessed: Aug. 10, 2022. [Online]. Available: http://proceedings.mlr.press/v32/le14.html?ref=https://githubhelp.com.

[13] A. Mandal, R. Chaki, S. Saha, K. Ghosh, A. Pal, and S. Ghosh, "Measuring similarity among legal court case documents," *ACM Int. Conf. Proceeding Ser.*, pp. 1–9, Nov. 2017, doi: 10.1145/3140107.3140119.

[14] C. Xia, T. He, W. Li, Z. Qin, and Z. Zou, "Similarity Analysis of Law Documents Based on Word2vec," *Proc. - Companion 19th IEEE Int. Conf. Softw. Qual. Reliab. Secur. QRS-C 2019*, pp. 354–357, Jul. 2019, doi: 10.1109/QRS-C.2019.00072.

[15] D. Thenmozhi, K. Kannan, C. A.-F. (Working Notes), and undefined 2017, "A text similarity approach for precedence retrieval from legal documents.," *ceur-ws.org*, Accessed: Aug. 10, 2022. [Online]. Available: http://ceur-ws.org/Vol-2036/T3-9.pdf.

[16] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," Accessed: Aug. 15, 2022. [Online]. Available: http://nlp.

[17] D. Cer *et al.*, "Universal Sentence Encoder," *AAAI*, pp. 16026–16028, Mar. 2018, doi: 10.48550/arxiv.1803.11175.

[18] W. H.Gomaa and A. A. Fahmy, "A Survey of Text Similarity Approaches," *Int. J. Comput. Appl.*, vol. 68, no. 13, pp. 13–18, Apr. 2013, doi: 10.5120/11638-7118.

[19] A. R. Lahitani, A. E. Permanasari, and N. A. Setiawan, "Cosine similarity to determine similarity measure: Study case in online essay assessment," *Proc. 2016 4th Int. Conf. Cyber IT Serv. Manag. CITSM 2016*, Sep. 2016, doi: 10.1109/CITSM.2016.7577578.

[20] X. Zhi, S. Yuexin, M. Jin, Z. Lujie, and D. Zijian, *Research on the Pearson correlation coefficient evaluation method of analog signal in the process of unit peak load regulation*. .

Appendix

Plagiarism Report[1]

AI-Enabled Automation Solution for Utilization Management in Healthcare Insurance

by Gaurav Karki

Submission date: 18-Aug-2022 10:35AM (UTC+0530)
Submission ID: 1883825378
File name: ilization_Management_in_Healthcare_Insurance_-_Gaurav_Karki.docx (908.38K)
Word count: 6023
Character count: 33983

[1] Turnitn report to be attached from the University.

AI-Enabled Automation Solution for Utilization Management in Healthcare Insurance

ORIGINALITY REPORT

8%	5%	3%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Queen's University of Belfast Student Paper	1%
2	Alfirna Rizqi Lahitani, Adhistya Erna Permanasari, Noor Akhmad Setiawan. "Cosine similarity to determine similarity measure: Study case in online essay assessment", 2016 4th International Conference on Cyber and IT Service Management, 2016 Publication	1%
3	Submitted to University of Strathclyde Student Paper	1%
4	Submitted to Liverpool John Moores University Student Paper	1%
5	hme-business.com Internet Source	1%
6	www.managedhealthcareexecutive.com Internet Source	1%
7	www.nh.gov	

	Internet Source	<1 %
8	www.cio.com Internet Source	<1 %
9	www.ama-assn.org Internet Source	<1 %
10	scholarbank.nus.edu.sg Internet Source	<1 %
11	Submitted to Lovely Professional University Student Paper	<1 %
12	3015elementcenter.com Internet Source	<1 %
13	Submitted to Nanyang Technological University Student Paper	<1 %
14	Submitted to University College London Student Paper	<1 %
15	data.dws.informatik.uni-mannheim.de Internet Source	<1 %
16	www.cookchp.org Internet Source	<1 %

Exclude quotes On Exclude matches < 10 words

Publications in Conference

Paper Submitted:

Gaurav Karki, Jay Bharateesh Simha, Rashmi Agarwal **"AI-Enabled Automation Solution for Utilization Management in Healthcare Insurance"**, EAICISML 2022 - EAI International Conference on Intelligent Systems and Machine Learning.

Submission Date: 22nd October 2022

AI-Enabled Automation Solution for Utilization Management in Healthcare Insurance

Gaurav Karki[1[0000-0003-0445-3261]] , Jay Bharateesh Simha[2[]] ,

Rashmi Agarwal[3[0000-0003-1778-7519]]

[1,2,3] REVA Academy for Corporate Excellence (RACE), REVA University
Bengaluru, India
[1]gauravk.ai01@race.reva.edu.in,
[2]jb.simha@reva.edu.in,
[3]rashmi.agarwal@reva.edu.in

Abstract. As businesses advance toward digitalization by automating an increasing number of procedures, unstructured forms of text in documents present new challenges. Most organizational data is unstructured, and this phenomenon is on the rise. Businesses like healthcare and insurance are embracing business process automation and making considerable progress along the entire value chain. Artificial intelligence (AI) algorithms that help in decision-making, connect information, interpret data, and apply the insights gained to rethink how to make better judgments are necessary for business process automation.

A healthcare procedure called Prior Authorization (PA) could be made better with the help of AI. PA is an essential administrative process that is a component of their utilization management systems, and as a condition of coverage, insurers require providers to obtain preapproval for the provision of a service or prescription. The processing of insurance claim documents can be facilitated using Natural Language Processing (NLP). This paper describes the migration of manual procedures to AI-based solutions in order to accelerate the process. The use of text similarity in systems for information retrieval, question-answering, and other purposes has attracted significant research. This paper suggests using a universal sentence encoder, a more focused strategy, to handle health insurance claims. By extracting text features, including semantic analysis with sentence embedding, the context of the document may be determined. The outcome would have a variety of possible advantages for members, providers, and insurers. AI models for the PA process are seen as promising due to their accuracy and speed of execution.

Keywords: Utilization Management, Prior Authorization, Healthcare, Insurance, Claim Processing, Deep Learning, Artificial Intelligence, Automation, Natural language processing.

1 INTRODUCTION

The expenses of the healthcare system have been spiraling out of control for years, and utilization management is a crucial method for insurers and providers to guarantee that adequate care is delivered in a cost-effective manner. Utilization management (UM) is the evaluation of medical care based on evidence-based criteria and health payer requirements. In addition to reducing costs, the goal is to give patients with the appropriate treatment, from a shorter duration of stay to enhanced release planning. Once insurers have defined their norms and guidelines, the utilization management process centers on controlling prior authorization filings via clinical and peer-to-peer evaluations [1].

The influence of UM processes on payer finances, case management, and health plan member and provider satisfaction is direct. Ineffective utilization control results in annoying care delays and higher operating expenses for the health insurance. However, conventional UM systems place member and provider satisfaction and cost control in opposition. More rules or more comprehensive review processes can reduce the high cost of medical care, but at the expense of member and provider satisfaction. Up until now, there has always been a compromise. Now, insurers have the chance to transform their UM processes through the application of automation and AI. By introducing an AI-based solution for utilization management, insurers can simultaneously enhance the member and provider experience while reducing operating expenses and medical costs [2].

PA is a primary process of UM. This is conducted before to or at the start of treatment on a particular scenario basis in an effort to eliminate needless services. The selected treatment should be considered provisional and is subject to modification in the future. PA is an examination of a patient's condition and suggested therapy. Its primary objective is to reduce unnecessary, ineffective, or redundant treatments. PA is utilized for regular and urgent referrals, but not for emergency room admissions. The review might take place either before or after admission, but always before treatment begins. In some situations, a physician's directions may not be followed, that could enrage both the medical staff and the patient.

This paper proposed solution tries to automate the PA procedure. Unlike basic Robotic Process Automation (RPA) systems that just automate individual aspects of the prior authorization process, artificial intelligence may fundamentally alter the way reviewers handle prior authorization requests. If a healthcare provider submits an authorization request, artificial intelligence evaluates this information to the medical necessity criteria to ensure that the patient receives the proper care. If all conditions are satisfied, the request may be automatically authorized, without the requirement for a utilization management reviewer to touch the previous authorization and a clinical evaluation. This can reduce approval times from weeks to hours.

In the field of text similarity, corpus-based techniques have solved the most challenging aspect of natural language processing by achieving human-competitive accuracy. However, later paper has shown that even a little difference in text structure or length can easily mislead the prediction. Term frequency inverse document frequency (TF-IDF) is a common approach presented by some that is believed to compensate for the inaccuracy introduced by the document's format and length, but at the expense of precision. The majority of previous text similarity techniques did not consider the embedding meaning of the words. When working with identical documents other than their wording, the ability of embedding meaning of words becomes useful.

2 LITERATURE REVIEW

PA in medical billing aids the healthcare organization in collecting the correct reimbursement for delivered services, hence lowering denials and subsequent follow-up. According to the results of the Prior Authorization Survey conducted by the American Medical Association (AMA) in December 2021 with the participation of 1,000 practicing physicians, the majority of physicians report an increase in the number of PAs necessary for prescription medications and medical services during the preceding years. In both instances as shown in Fig. 1, the proportion of physicians reporting this rise ranged from 84% to 84% [3].

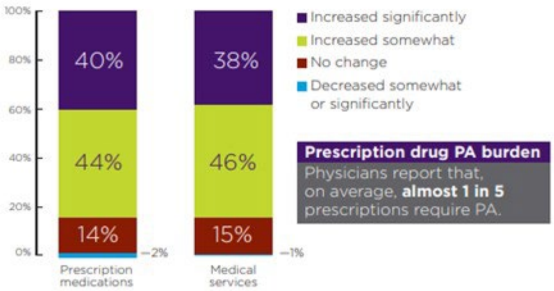

Fig. 1. AMA Prior Authorization Physician Survey [3]

Bag-Of-Words (BOW) [4] and Term Frequency-Inverse Document Frequency (TF-IDF) [5] models have been widely used for text encoding by conventional machine learning algorithms in various text analytics domains, such as the legal sector. Kumar et al. observed that the BOW and TF-IDF model yields better results for recognizing the similarity of legal judgments since just the similarity of the legal phrases, as opposed to all the terms in the dataset, are evaluated [6].

Mandal et al. explored multiple advanced vector representation models for legal document such as Latent Dirichlet Allocation [7] and word embedding along with the TF-IDF model. Word embedding techniques such as Word2vec [8] and Doc2vec [9], that may better capture the semantics of documents, achieved 69% Pearson correlation coefficients in finding similarities among Indian Supreme Court cases, according to the paper [10].

Word embedding representation has been used in the legal field for a number of years. However, the results reported in the reviewed existing literature [10], [11], [12] provide a marginal or non-existent improvement over simple BOW models. This paper investigates the pre-trained, domain-specific deep learning model utilized for text embedding. After exploring for comparable instances, the newly installed embedding yielded superior results.

3 METHODOLOGY

The method proposes in this work offers an approach for screening PA cases based on local coverage determinations, that are decisions made by a Medicare Administrative Contractor (MAC) whether to approve a case or not. This is an NLP-based model to help insurers by screening the PA cases that give the most similar information and ranking them based on the similarity score in comparison with the Local Coverage Determinations (LCDs) criterions. The Proposed solution consists of the following components as shown in Fig. 2.

1. PA case
2. Requested process id and Healthcare Common Procedure Coding System (HCPCS) code
3. LCDs Criterions
4. Lexicons
5. Text Similarity Check

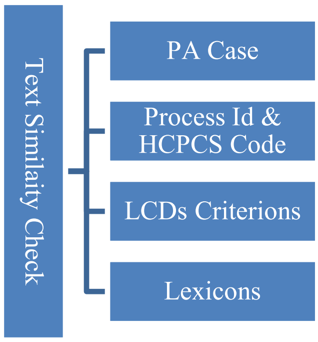

Fig. 2. The Proposed Solution's Components

PA cases are collected through any channel between Insurer and provider, including Electronic Medical Record (EMR). Then, getting prepared three JSON files containing all the necessary information to be utilized as input for the text similarity model.

- The first file contains the rules from the LCD criteria used to approve the procedure in relation to the HCPCS code.
- The second file contains the lexicons associated with each rule; these lexicons aid in the identification of required information from PA cases.

- The third file contains both the requested process id and the HCPCS code needed to verify the eligibility of PA case.

The model retrieves text from the lexicon file and PA case and then executes all of the steps outlined in the following Fig. 3.

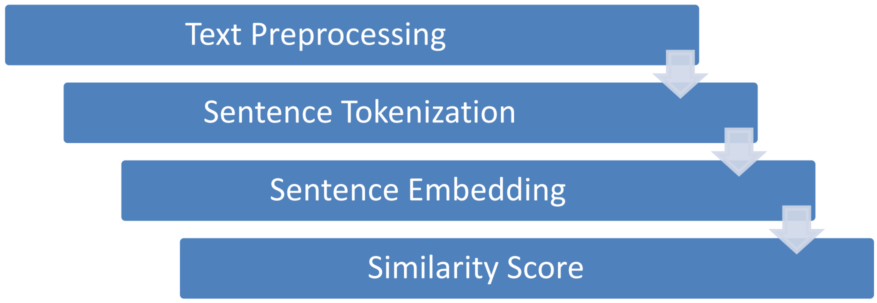

Fig. 3. Text Analysis Workflow

4 SOFTWARE DESIGN

As depicted in Fig. 4, the suggested AI-enabled solution consists primarily of three phases. First, the provider submits a PA request, then the PA automation engine examines the case and give the necessary information from the text to assist Registered Nurses (RNs) in making an approval or denial determination. Now examine the automation engine's underlying framework.

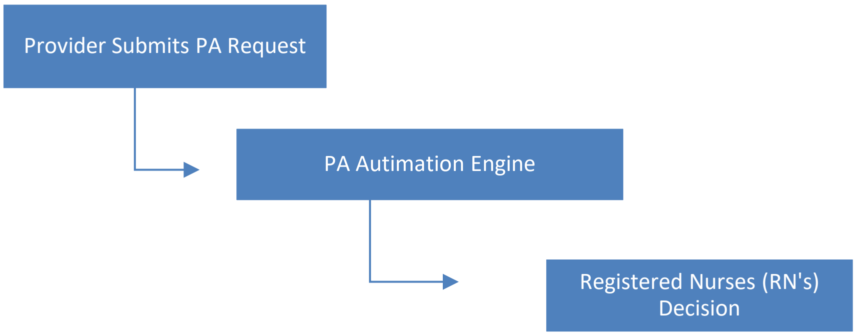

Fig. 4. Proposed AI-Enabled PA Process

There are three stages involved in the operation of an automation engine as shown in Fig. 5. The first stage involves process verification; the second stage identifies rules associated with process id and HCPCS code, as a result, finalizes the lexicons for each rule. and the third stage examines textual similarities. If the process verification phase is successful, the subsequent phases are enabled; if it is failed, a case of rejection is possible. Each step is elaborated addressed separately.

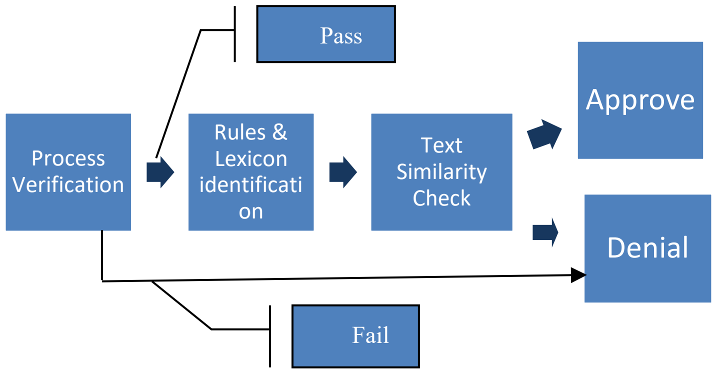

Fig. 5. Automation Engine Process

5 IMPLEMENTATION

In an effort to create an AI-enabled solution. The first and possibly most crucial phase is data preparation, because the quality of the solution depends on the quality of the data used. On the payer side, data is collected in accordance with rules. All essential information is kept in JSON format in a configuration folder containing data collected from insurers in accordance with their requirements. The AI-enabled solution for PA requires two components for implementation. The provider provides PA cases in text format with HCPCS codes. Once a PA case containing the requested HCPCS code is received, the automation engine is engaged. According to the design, the automation engine consists of three stages.

The first stage is process verification and this procedure begins with the input of the HCPCS code. Therefore, the automation engine employs a rule-based method to cross-check with json files containing the process id and corresponding codes. If this eligibility check is successful, the procedure advances to the subsequent level; otherwise, the case is denied.

The second stage is rules and lexicon identification. HCPCS code facilitates the identification of rules using a rule-based methodology. Consequently, lexicon is identified relative to each rule using the same method.

The third and last important stage is text similarity. This stage is the central concept of AI utilization in this paper. Textual analysis comes into play here.; it aids with text comprehension. Text is provided as input at this stage, after that text preparation occurs. Then, the sentence is tokenized so that the model can convert it into a vector, that is the sentence embedding process. The identical procedure of sentence embedding was used to the lexicon identified in the previous step. Now, after obtaining two sets of vectors, one from lexicons and the other from text, cosine similarity is used to generate similarity metrics. After receiving ranking evidence, RNs can make a decision based on this evidence.

Sentence embedding is performed to generate a dictionary of pre-processed sentences and their corresponding vectors. Vectors are derived using a model. This model is founded on NLP principles. In this paper, two NLP strategies were utilized to determine the most effective method for achieving our objective.

5.1 TF-IDF

In this approach, the frequency of words is rescaled according to their overall frequency in all texts, which penalizes common, ubiquitous words like "the" that appear often throughout all texts. TF-IDF measures how crucial a specific term is to the overall meaning of a text. Multiplying two separate metrics yields a document word's TF-IDF. The Term Frequency (TF) of a document's words. There are numerous methods for calculating this frequency, the simplest of that is a simple count of the occurrences of a word in a document. Then, there are further methods for adjusting the frequency. For instance, as Equation (1) describes, by dividing the raw count of occurrences of a word by the document's length or by the raw frequency of the document's most frequent word.

$$TF(i,j)= n(i,j) / \Sigma\, n(i,j) \tag{1}$$

Where,
$n(i,j)$ = number of times nth word occurred in a document
$\Sigma\, n(i,j)$ = total number of words in a document.

The inverse document frequency (IDF) of a given word across a collection of documents. This reflects the frequency of a word in the entire document set. The closer a term is to 0, the more frequent it is. This metric can be determined by dividing the total number of documents by the number of documents containing a specific word. This metric can then be calculated using the logarithm.

Therefore, this number approaches 0 if the term is prevalent and appears in several documents. Alternatively, it approaches 1 Multiplying these two numbers yields the TF-IDF score of each word in a document. The higher the score, the more pertinent the word is to the document. In mathematical terms, the TF-IDF score is calculated according to Equation (2).

$$IDF=1+\log(N/dN) \tag{2}$$

Where,
N =Total number of documents in the dataset
dN =total number of documents in that nth word occur

5.2 UNIVERSAL SENTENCE ENCODER

A significant amount of work is expended in machine learning research to convert data into vectors. Word2vec and Glove [13] accomplish this by turning a word into a vector. Therefore, the vector corresponding to "cat" will be closer to "dog" than to "eagle." While embedding a sentence with its words, however, the complete sentence's context must be captured in that vector. The "Universal Sentence Encoder" comes into play at this point.

The embedding generated by the Universal Sentence Encoder [14] model especially transfer learning to the NLP tasks. It is trained on a number of data sources in order to acquire skills for a vast array of tasks. The sources include Wiki, web media, online question-and-answer pages, and forums. The input is variable-length English text, while the outcome is a 512-dimensional vector.

Typically, sentence embedding was derived by averaging the embedding of all the words in the phrase; however, this method had limitations and was unsuitable for detecting the true semantic meaning of a sentence. The Universal Sentence Encoder makes sentence-level embedding effortless. It is available in two variants, one trained with the Transformer encoder and the other with the Deep Averaging Network (DAN). In terms of computer resource requirements and accuracy, there is a trade-off between the two. While the one with the Transformer encoder is more precise, it requires more computation. The variant with DAN encoding is computationally less expensive and slightly less precise. This paper utilizes the transformer encoder variant.

6 ANALYSIS AND RESULTS

In this paper, analysis is done by STS Benchmark. This Benchmark offers an empirical assessment of the degree to which similarity ratings obtained by sentence embedding correspond to human judgments. The benchmark necessitates that systems produce similarity scores for an assortment of sentence pairs. The Pearson correlation coefficient is then applied to compare the quality of algorithm similarity scores to human judgments. The statistical method of the Pearson correlation coefficient [15] is commonly used in economics for purposes like trend analysis and classification. Other potential domains of use have been discussed in recent years' literature. Using it, one can determine if strongly two variables are related to one another along a linear axis. And p-value is calculated to check statistically significance. If the correlation coefficient were indeed zero, then the current result would have been seen with a probability equal to the P-value (null hypothesis). A correlation coefficient is considered statistically significant if its associated probability is less than 5%.

6.1 TF-IDF

Pearson correlation coefficient = 0.2340
p-value = 1.015e-19

6.2 UNIVERSAL SENTENCE ENCODER

Pearson correlation coefficient = 0.83
p-value = 0.0

And testing and validation purpose only few rules have been considered that are Rule "A", Rule "B" and Rule "C".

- Rule "A" says one or more mobility-related activities of daily living, are severely hindered because of the beneficiary's mobility limitation.
- Rule "B" says a properly adjusted cane or walker would not help the beneficiary with his/her mobility issues to an acceptable degree.
- Rule "C" says there is insufficient upper extremity function for the beneficiary to propel a properly equipped manual wheelchair indoors.

Results that are from the TF-IDF technique are mentioned in the Table 1.

Table 1. Results from the TF-IDF Technique

Rule Name	Total no. of Matches	Top matching sentence from PA text	Highest Cosine Score
A	5	he limited in his ability to participate in all mobility related activities of daily living in the home setting	0.491431
B	6	he la unable to safely or effectively use cane or walker for the distance needed in the home due to fatigue joint pain and numbness	0.2933
C	9	he is unable to self-propel an optimally configures manual wheelchair due to upper extremity weakness and arthritic hand pain	0.3841

Results that are from the Universal Sentence Encoder Technique are mentioned in the Table 2.

Table 2. Results from the Universal Sentence Encoder Technique

Rule Name	Total no. of Matches	Top matching sentence from PA text	Highest Cosine Score
A	7	he limited in his ability to participate in all mobility related activities of daily living in the home setting	0.72
B	5	he la unable to safely or effectively use cane or walker for the distance needed in the home due to fatigue joint pain and numbness	0.70
C	8	he is unable to self-propel an optimally configures manual wheelchair due to upper extremity weakness and arthritic hand pain	0.71

So, the result from the both technique says Universal Sentence Encoder has a high correlation coefficient, p-value and cosine score compared to TF_IDF.

7 CONCLUSION

The facts presented in the previous section make it plainly clear that Universal Sentence Encoder's tactics are superior to those of TF IDF. Using the Pearson correlation coefficient, the assessment standard allows for differentiation between the various methodologies. The similarity visualization generates outputs with varying degrees of color intensity that are quite similar. This conclusion is based on the information gathered for the purpose of this paper. On the other hand, the evaluation benchmark makes it obvious that Universal Sentence Encoder approaches can still tackle the challenge even if the difficulty of the text increases. Even though this paper is conducted on PA cases, that are part of the healthcare industry, complexity management is always a concern. This methodology aids in reaching the objectives of the paper. Access to necessary care for patients is frequently delayed as a result of prior authorizations, that may drive patients to abandon their treatment due to the waiting period or other complications related with prior authorization. This work provides a viable method for resolving the issue, as it proposes a method for streamlining AI that minimizes treatment delays and disruptions by reducing the requirement for prior approval. This solution, that is an integral part of the End-to-End Prior Authorization process, eliminates human work that is time-consuming and prone to error. Therefore, the pre authorization team can maximize the health system's capacity to provide faster and better care. Patients, healthcare providers, and insurers, as well as any other parties engaged in the process, can all benefit from an efficient utilization management program. These are the adverbial complements for each:

- Patients gain from decreased treatment costs, more treatment efficacy, and fewer refused claims.
- Fewer denied claims, reduced costs, more effective treatments, improved data, and more efficient resource utilization are all beneficial to the health care industry.

References

1. T. M. Wickizer and D. Lessler, "Utilization management: Issues, effects, and future prospects," Annu. Rev. Public Health, vol. 23, pp. 233–254, 2002, doi: 10.1146/ANNUREV.PUBLHEALTH.23.100901.140529..
2. "AI ushers in next-gen prior authorization in healthcare | McKinsey | McKinsey." https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/ai-ushers-in-next-gen-prior-authorization-in-healthcare (accessed Aug. 10, 2022).
3. A. Medical Association, "Prior Authorization Physician Survey Update | AMA," 2022, Accessed: Aug. 10, 2022. [Online]. Available: https://www.ama-assn.org/system/files/prior-.
4. E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," Decis. Support Syst., vol. 50, no. 3, pp. 559–569, 2011, doi: 10.1016/J.DSS.2010.08.006.
5. J. Lam, Y. Chen, F. Zulkernine, and S. Dahan, "Detection of Similar Legal Cases on Personal Injury," IEEE Int. Conf. Data Min. Work. ICDMW, vol. 2021-December, pp. 639–646, 2021, doi: 10.1109/ICDMW53433.2021.00084.
6. M. Kumar, R. Ghani, and Z. S. Mei, "Data mining to predict and prevent errors in health insurance claims processing," Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 65–73, 2010, doi: 10.1145/1835804.1835816.
7. D. Blei, A. Ng, M. J.-J. of machine L. research, and undefined 2003, "Latent dirichlet allocation," jmlr.org, vol. 3, pp. 993–1022, 2003, Accessed: Aug. 10, 2022. [Online]. Available: https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf?ref=https://githubhelp.com.
8. T. Mikolov, K. Chen, G. Corrado, J. D. preprint arXiv:1301.3781, and undefined 2013, "Efficient estimation of word representations in vector space," arxiv.org, Accessed: Aug. 10, 2022. [Online]. Available: https://arxiv.org/abs/1301.3781.

9. Q. Le, T. M.-I. conference on machine, and undefined 2014, "Distributed representations of sentences and documents," proceedings.mlr.press, Accessed: Aug. 10, 2022. [Online]. Available: http://proceedings.mlr.press/v32/le14.html?ref=https://githubhelp.com.
10. A. Mandal, R. Chaki, S. Saha, K. Ghosh, A. Pal, and S. Ghosh, "Measuring similarity among legal court case documents," ACM Int. Conf. Proceeding Ser., pp. 1–9, Nov. 2017, doi: 10.1145/3140107.3140119.
11. C. Xia, T. He, W. Li, Z. Qin, and Z. Zou, "Similarity Analysis of Law Documents Based on Word2vec," Proc. - Companion 19th IEEE Int. Conf. Softw. Qual. Reliab. Secur. QRS-C 2019, pp. 354–357, Jul. 2019, doi: 10.1109/QRS-C.2019.00072.
12. D. Thenmozhi, K. Kannan, C. A.-F. (Working Notes), and undefined 2017, "A text similarity approach for precedence retrieval from legal documents.," ceur-ws.org, Accessed: Aug. 10, 2022. [Online]. Available: http://ceur-ws.org/Vol-2036/T3-9.pdf.
13. J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," Accessed: Aug. 15, 2022. [Online]. Available: http://nlp.
14. D. Cer et al., "Universal Sentence Encoder," AAAI, pp. 16026–16028, Mar. 2018, doi: 10.48550/arxiv.1803.11175.
15. X. Zhi, S. Yuexin, M. Jin, Z. Lujie, and D. Zijian, Research on the Pearson correlation coefficient evaluation method of analog signal in the process of unit peak load regulation.