**Enron Investigation**

**A Project Work Synopsis**

*Submitted in the partial fulfillment for the award of the degree of*

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE WITH SPECIALIZATION IN**

**ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**Submitted by:**

20bcs6766 - Yamini

20bcs6746 - Saumya Dua

20bcs6788 - Mayank

**Under the Supervision of:**

**<Project Supervisor Name>**



**CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,**

**PUNJAB**

**September, 2022**

# Abstract

The Enron Corporation scandal of 2001 was one of the largest and most devastating corporate fraud cases in history, resulting in the company's bankruptcy and the downfall of its top executives. This project aims to investigate the events leading up to the scandal and analyze the factors that contributed to its occurrence.

Our team will be working on Enron Email data to get insights. Building a web application which provides interactive visualization of the data using Machine Learning is the goal.

Through a comprehensive analysis of the Enron scandal, this project aims to shed light on the systemic issues that contribute to corporate fraud and to identify lessons that can be learned to prevent similar events from occurring in the future.

# Table of Contents

| Title Page | i |
| --- | --- |
| Abstract | ii |
| 1.    Introduction |  |
| 1.1           Problem Definition |  |
| 1.2 Project Overview |  |
| 1.3 Hardware Specification |  |
| 1.4 Software Specification |  |
| 2.    Literature Survey |  |
| 2.1 Existing System |  |
| 2.2 Proposed System |  |
| 2.3 Literature Review Summary |  |
| 3.    Problem Formulation |  |
| 4.    Research Objective |  |
| 5.    Methodologies |  |
| 6.    Experimental Setup |  |
| 7.    Conclusion |  |
| 8.    Tentative Chapter Plan for the proposed work |  |
| 9.    Reference |  |

# 1. INTRODUCTION

## 1.1 Problem Definition

## 1.2 Problem Overview

## 1.3 Hardware Specification

1. Computer: A high-performance computer with a multi-core CPU, at least 8GB of RAM.

2. Storage: Sufficient storage to store the Enron data.

3. Display: A high-resolution display to visualize and interact with the data.

4. Web server: A web server to host the web application.

5. Cloud service: A cloud service to deploy the web application, such as Heroku

## 1.4 Software Specification

1. Python: Python is a popular programming language for machine learning and web development.

2. Machine learning libraries: There are several machine learning libraries available for Python, including scikit-learn, TensorFlow, and Keras.

3. Web development frameworks: There are several web development frameworks available for Python, including Flask.

4. Text editors: A text editor is necessary for writing and editing code. Such as Visual Studio Code, and Sublime Text.

5. Database software: Database is required to store and manage the Enron data. Some popular options for Python include PostgreSQL.

# 2. LITERATURE SURVEY

## 2.1 Existing System

Enron scandal is just a case study as of now. Multiple papers and documents have been brought out to understand the scandal

## 2.2 Proposed System

Building web application to understand this scandal more deeply, get deeper insights through enron email dataset will be a step forward towards the understanding, also providing accessibility to general users at the same time.

## 2.3 Literature Review Summary (Minimum 7 articles should refer)

| **Year and**  **Citation** | **Article/ Author** | **Tools/ Software** | **Technique** | **Source** | **Evaluation Parameter** |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

.

# 3. PROBLEM FORMULATION

The Enron scandal is an important case study in corporate governance, accounting ethics, and financial fraud. It serves as a cautionary tale for businesses, regulators, and investors alike, highlighting the dangers of unethical behavior, tax regulation, and the need for transparency and accountability in the corporate world.

By analyzing the Enron scandal, we can identify the systemic issues that contributed to the fraud and develop strategies to prevent similar events from occurring in the future. It also highlights the importance of effective corporate governance, strong internal controls, and independent oversight.

Moreover, the Enron scandal had significant social and economic implications, leading to job losses, investor losses, and a loss of public trust in the financial system. By studying the scandal, we can understand the human impact of corporate fraud and the importance of ethical business practices in promoting long-term economic growth and social stability.

Analyzing the Enron scandal is important because it helps us learn from past mistakes, prevent future corporate fraud, and promote ethical behavior in the corporate world, benefiting both businesses and society as a whole.

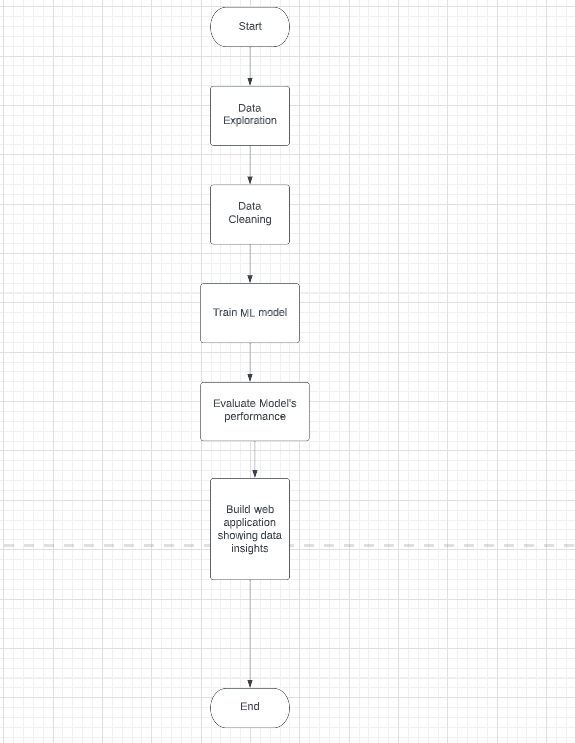
Building the Enron Scandal project is a necessity as it will help provide the general public an easy understanding and insights to the scandal would benefit the geeks and society.

# 4. OBJECTIVES

* **Collect and preprocess data** - Start by collecting relevant data related to the Enron scandal, such as financial statements, email communications, and news articles. Preprocess the data by cleaning, transforming, and formatting it for analysis.
* **Define ML tasks** - Define the ML tasks that you want to perform on the data, such as classification, clustering, or regression. For example, you could use classification to identify Enron employees involved in the scandal or use clustering to group similar email communications.
* **Train ML models** - Train ML models using various algorithms and techniques, such as decision trees, neural networks, and deep learning. Use cross-validation and hyperparameter tuning to optimize the models' performance.
* **Evaluate model performance** - Evaluate the performance of the ML models using metrics such as accuracy, precision, recall, and F1 score. Use visualizations such as confusion matrices, ROC curves, and precision-recall curves to interpret the model's performance.
* **Develop Web Application** - Develop a web application that showcases the ML models and provides interactive visualizations of the data. Use frameworks such as Flask or Django for the web application and deploy it on a cloud service such as Heroku.
* **Create a presentation**: Create a presentation that summarizes the project's objectives, methodology, and results. Include a demo of the web application and explain how the ML models can be used to gain insights into the Enron scandal.

# 5. METHODOLOGY

1. **Data Exploration** - Exploratory Data Analysis (EDA) to analyze and explore email dataset
2. **Data Cleaning** - Remove special characters, stopwords, words with similar root meaning, trim email duplicates
3. **ML Model** - Using algorithm that maps words within a document in a given corpus to a vector of floats
4. **Machine Learning Algorithm** - Use clustering to group emails with same labels together and evaluating model’s performance
5. **Web Application** - Building web application to provide data insights and deploying it on cloud



Following technologies will be used in implementing Enron Scandal project

* Python
* ML Libraries - scikit learn, tensorflow, keras
* Flask
* Postgres SQL
* Heroku

# 6. EXPERIMENTAL SETUP

# 7. CONCLUSION

Data is the new fuel and Enron project will use it to bring amazing insights into the well known scandal of all times.

Building a Web Application out of this case study using ML models will help users get insights into the Enron Scandal through visualizations, making people more aware.

Any user will be able to understand what happened to Enron.

## 8. TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

**CHAPTER 1: INTRODUCTION**

**Week 1-2**

**CHAPTER 2: LITERATURE REVIEW**

**Week 1-2**

**CHAPTER 3: OBJECTIVE**

**week 1-2**

**CHAPTER 4: METHODOLOGIES**

**week 3-4**

**CHAPTER 5: EXPERIMENTAL SETUP**

**week 3-4**

**CHAPTER 6: CONCLUSION AND FUTURE SCOPE**

**week 3-4**

## REFERENCES

[1] <https://www.investopedia.com/updates/enron-scandal-summary/>

<https://en.wikipedia.org/wiki/Enron_scandal>

<https://www.britannica.com/event/Enron-scandal>

<https://www.wallstreetmojo.com/enron-scandal/>

<https://www.journalofaccountancy.com/issues/2002/apr/theriseandfallofenron.html>