REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

REVA Academy for Corporate Excellence (RACE)

Reducing Business Compliance Risk with help of Analytics

Lalit Aggarwal

SRN: R19DM003

Date: 20th August 2022

PGDM/MBA in Business Analytics

Capstone Project Presentation
Year: I

race.reva.edu.in

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Agenda

01 Introduction
Back Ground | Current status | Why this study

02 Literature Review
Seminal works | Summary | Research Gap

03 Problem Statement
Business Problem | Analytics Solution

04 Project Objectives
Primary & Secondary Objectives | Expected Outcome

05 Project Methodology
Conceptual Framework | Research Design

06 Business Understanding
Business Context | Monetary Impact

07 Data Understanding
Data Collection | Variables

08 Data Preparation
Pre-processing | Process | Techniques

09 Descriptive Analytics
Univariate | Bivariate | Hypothesis

10 Modeling
Machine Learning | Model Evaluation | Insights

11 Model Deployment
Applications | Demo

12 Suggestions and Conclusions
Insights | Next Step | Future Scope

13 Annexure
References | Publications | Plagiarism Score

Introduction

Background | Current status | Why this study

The broad use of the term "compliance risk" by financial services companies, and banks, is to describe the risks associated with the way organizations and their employees harm customers or negatively affect market stability.

Compliance risk is an organization's potential exposure to legal penalties, financial loss and material loss, resulting from its failure to act in accordance with industry laws and regulations, internal policies or prescribed best practices. Compliance risk is also known as integrity risk.

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Literature Review

Sr.No.	Title	Author	Detailed Study
1	Sales Culture and Misconduct in the Financial Services Industry:An Analysis of Cross-Selling Practices	Francesco De Pascalis (2018)	In this paper author highlited the ilegal practice used in **the cross-selling in the banking sector** and emphasizes intervention in the conduct, culture, and governance framework of financial institutions.
2	Misconduct Risk, Culture, and Supervision	James Hennessy et, al. (December 2017)	In this paper author highlited the wrongdoing in financial institutions and the **role of banking rules and managers** to mitigate risk by analyzing risk administration, interior controls, and governance.
3	Fraudulent Financial Reporting and the Consequences for Employees	Jung Ho ChoiBrandon Gipper (March 2021)	Author studies the impact of **dishonest reports against an institution** on their workers. If dishonest financial reporting happens against any organization, it may affect Employees' wages and their turnover before and after the incident.
4	Linking employee misbehavior to consumer satisfaction	Namasivayam, K. (2006)	In this paper author discussed linking between **employee misconduct and consumer satisfaction.** With mis handling of customer detriment their interest in doing business with the organization.
5	The Market for Financial Adviser Misconduct	Mark Egan, G. M. (2019)	Mark Egan stresses on the financial advisor misconducts, as per author around **7 to 15% advisor** are usually involved in these type of misconduct and they do it repeatedly. It's a common practice when some firms retrench them due to compliance some other firms rehire them.

Problem Statement

Business Problem | Analytics Solution

The purpose is to predict the compliance violators based on their past data in the organization. So organization can take a corrective measure in advance.

In this project we analyzed the employees past data and with the help of different Machine Learning algorithms and developed an efficient statistical models, which can give the probability of an employee might deviate from organization compliance policy in future. So by identifying such employees an organization can take a preventive action to reduces the financial and reputational losses that could happen due to violation of market standards.

Project Objectives

Primary & Secondary Objectives | Expected Outcome

The objective of this study is to categorize the employee of the organization in a separate section based on their business compliance risk. So the employer can take a preventive action by providing proper training or make them aware about the consequence of the bad compliance to the business . That way, an organization can avoid unnecessary monetary or reputational loss.

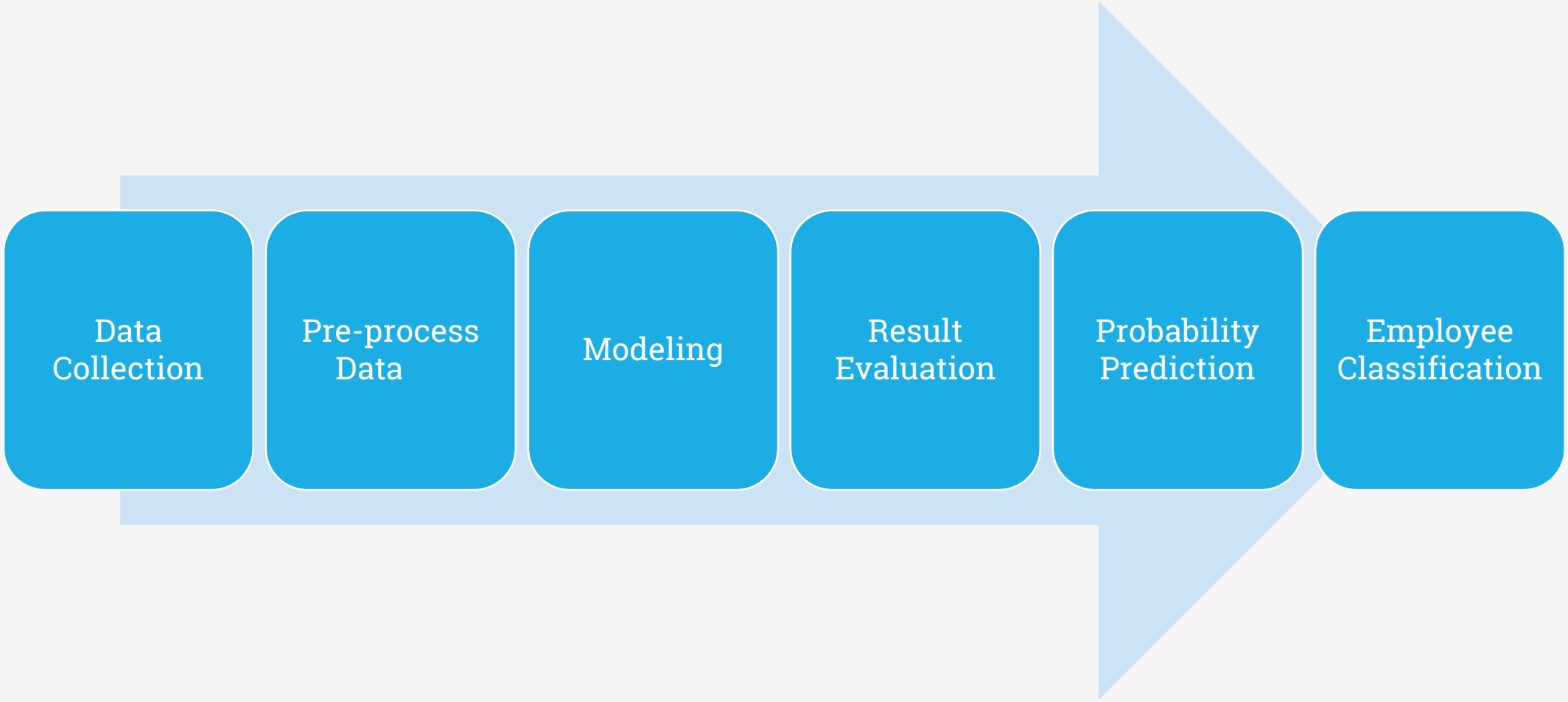
In this project we divided all the employees in to three categories based on the violation risks :

1. Low Risk
2. Medium Risk
3. High Risk

So the organization could concentrate to the employee who are laying in High Risk bracket.

Project Methodology

Conceptual Framework | Research Design



Business Understanding

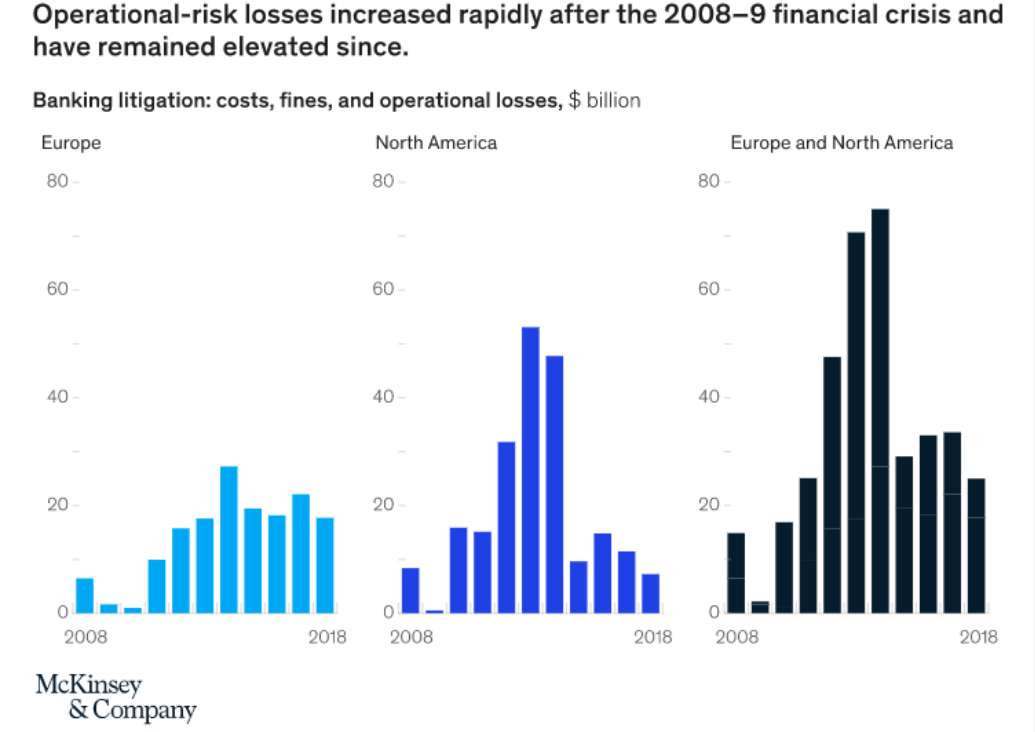
Business Impact | Challenges | Monetary Impact

There has been a substantial financial impact on organizations due to conduct-related regulatory action and it can all arise from a person's activities. The current report from the **Fixed Payment, Currencies and Commodities Markets Standards Board (FMSB) gauges banks have paid some $375 billion** in lead damages in the most recent long term.

Associations that fail to align lead hazards face administrative activity, damages, and reputational harm, which can hurt a business for quite a long time past the function.

e.g. Karvy, Satyam, DHFL,….

1. ABG Shipyard Bank Scam (Rs 22, 848 Cr, 2012-17)
2. Nirav Modi PNB Bank Fraud (Rs14,000 Cr, 2018)
3. Karvy Stock Broking Limited (Rs 2800 Cr, 2016-19)



Data Understanding

Data is obtained from internet which has 20 Independent and 1 dependent variable.

It's a data of bank employees which includes

1. Customer related data
2. Product related data
3. Employee related data

Target is the Dependent variable has two values

0 : Non Violator

1 : Violators

Independent variable

Account Usage data

0- Single Login
1- Multiple Login

Customer segment

1	Lowest net worth customer
2	Lower medium net worth customer
3	Medium net worth customer
4	Upper medium net worth customer
5	Highest net worth customer

Misaligned Incentives

1	lower range incentives
2	lower medium range incentives
3	medium range incentives
4	higher medium range incentives
5	higher range incentives

Employee performance

1-lowest
5-highest

Customer Feedback

1-lowest/completely unhappy
10-highest/extermely happy

compliance

1-least compliant to companies policies/ guidelines
5-highly compliant

Data Preparation

Pre-processing | Techniques

```
i,j in df.iterrows():
if (j['balance'] <= 1000000) | (j['income_levels'] <= 800000) | (j['tenure_with_bank'] <= 5) & (j['credit_score'] <= 500) | (j['
    j['Customer_segment'] = 1
elif (1000000 > j['balance'] <= 2000000) | (800000 > j['income_levels'] <= 1200000) | (5 > j['tenure_with_bank'] <= 10) | (500 >
    j['Customer_segment'] = 2
elif (2000000 > j['balance'] <= 4000000) | (1200000 > j['income_levels'] <= 2000000) | (10 > j['tenure_with_bank'] <= 20) | (600
    j['Customer_segment'] = 3
elif (4000000 > j['balance'] <= 6000000) | (2000000 > j['income_levels'] <= 3000000) | (20 > j['tenure_with_bank'] <= 30) | (750
    j['Customer_segment'] = 4
elif (j['balance'] > 6000000) | (j['income_levels'] > 3000000) | (j['tenure_with_bank'] > 30) | (j['credit_score'] > 820) | (j['
    j['Customer_segment'] = 5

'Customer_segment'].value_counts()
```

Customer segment created based on different customer related parameters.

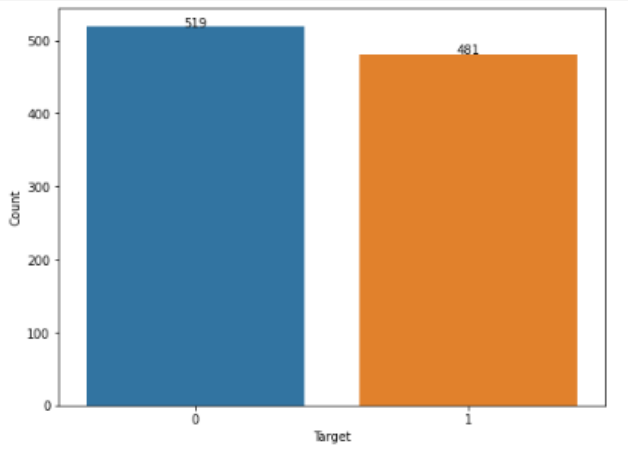
```
for m,n in df.iterrows():
    if ((n['points_of_improvement'] == 10) & (26 > n['retail_acc_setup_errrate'] <= 30) & (65 > n['avg_time_to_close_issues'] <=
        n['Employee_performance'] = 1
    elif ((8 >= n['points_of_improvement'] <= 9) & (19 > n['retail_acc_setup_errrate'] <= 26) & (55 > n['avg_time_to_close_issues
        n['Employee_performance'] = 2
    elif ((6 >= n['points_of_improvement'] <= 7) | (16 >= n['retail_acc_setup_errrate'] <= 20) | (45 > n['avg_time_to_close_issu
        n['Employee_performance'] = 3
    elif ((4 >= n['points_of_improvement'] <= 5) | (10 > n['retail_acc_setup_errrate'] <= 15) | (30 > n['avg_time_to_close_issues
        n['Employee_performance'] = 4
    elif ((n['points_of_improvement'] <= 3) | (2 >= n['retail_acc_setup_errrate'] <= 10) | (n['avg_time_to_close_issues'] <= 30)
        n['Employee_performance'] = 5

df['Employee_performance'].value_counts()
```

Employee performance segmented based on employee data

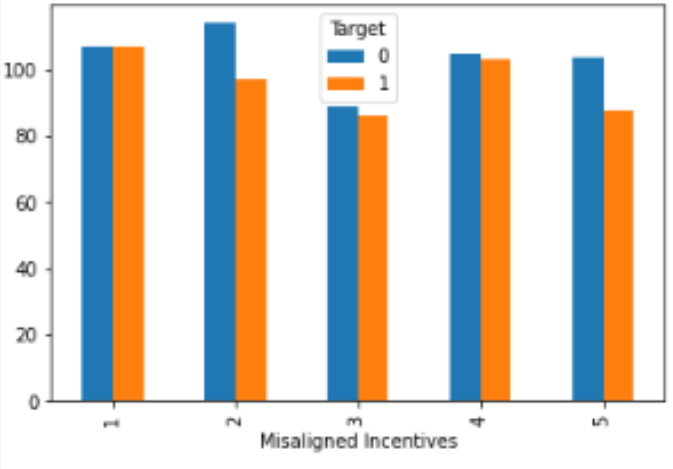
```
#   Column
---  ------
0   Emp_ID
1   Sales
2   Account_usage_data
3   Customer_segment
4   Misaligned_incentives
5   Employee_performance
6   Customer_feedback
7   points_of_improvement
8   compliance
9   retail_acc_setup_errrate
10  avg_time_to_close_issues
11  Product_performance
```

Descriptive Analytics

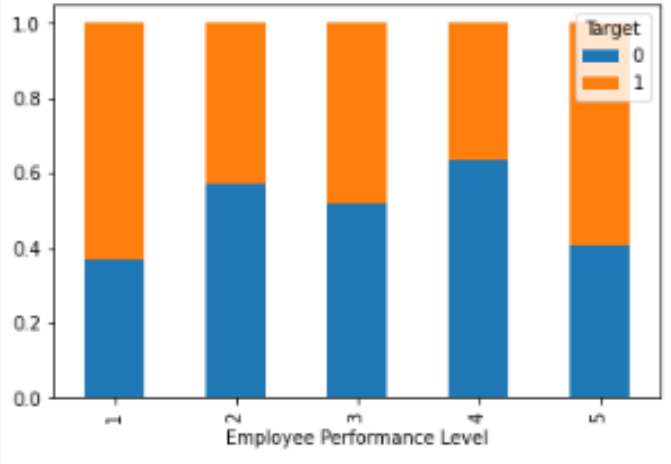
Multivariate Analysis | Hypothesis



Depended Feature :Target
Data is looking very balanced.



Misaligned Incentive vs Target
Lower incentive have balanced violations both are high.



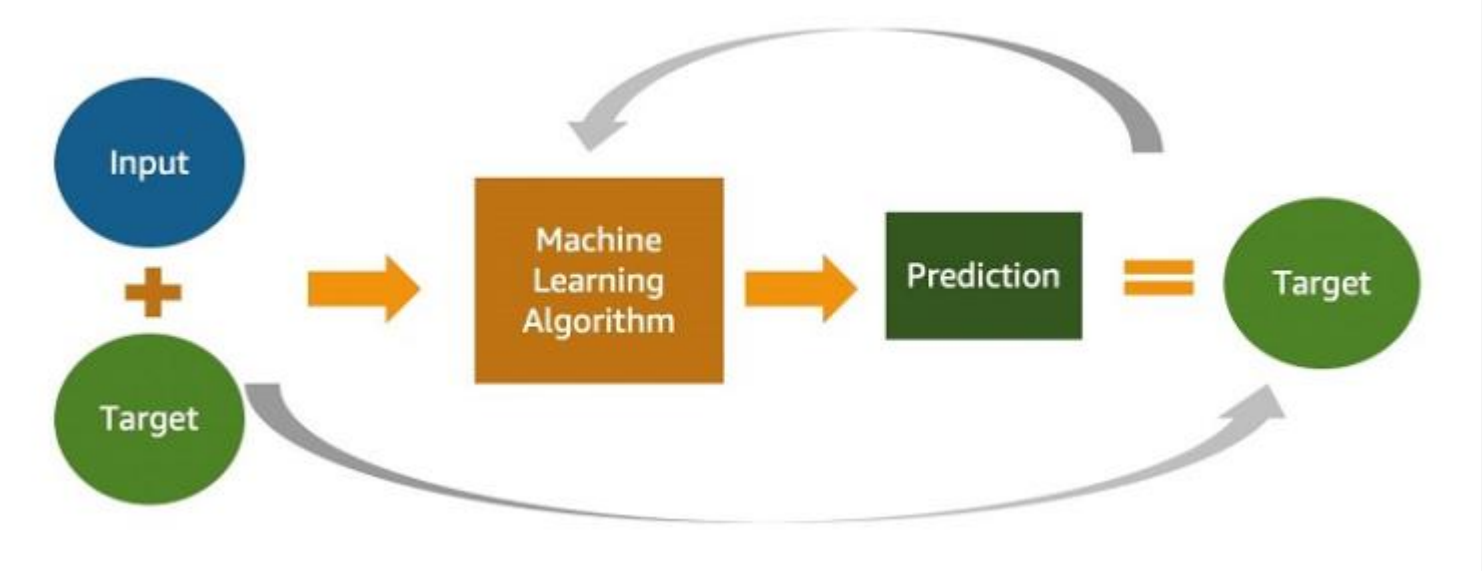
Employee Performance vs Target
Segment 1 and 5 have more violators

Modeling

Modeling Techniques | Modeling Process | Model Building

In this we've used following Machine learning techniques:

1. Logistic regression
2. K- Nearest Neighbors
3. Naïve Bayes
4. Decision Tree
5. Gradient Boosting



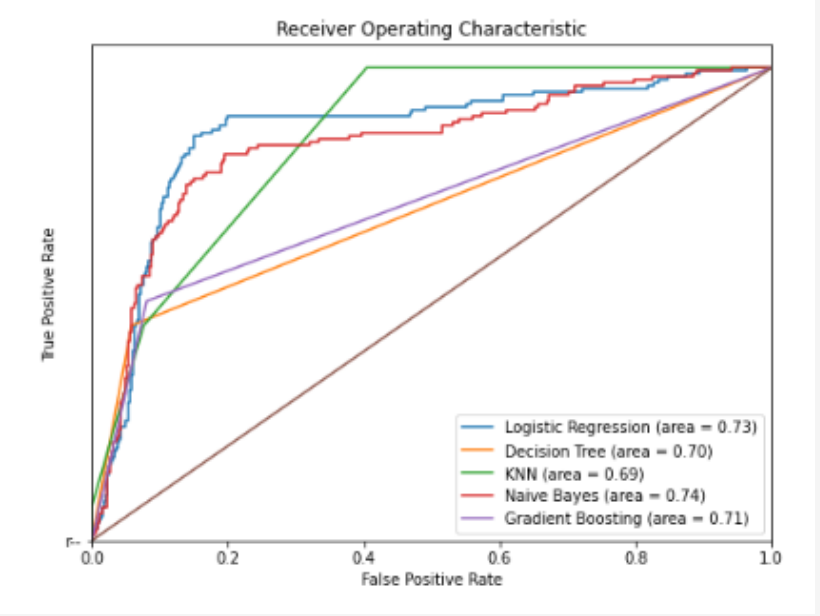
Model Evaluation

Results | Interpretation | Insights

Name of Algorithims	Accuracy	Precision	Recall	F1-Score
Logistic Regression	**80**	**78**	**77**	**77.5**
Decision Tree	77	68	70	68.98
KNN	69	66	63	64.47
Naïve Bayes	77.6	72	76	74
Gradient Boosting	78	70	72	70.98

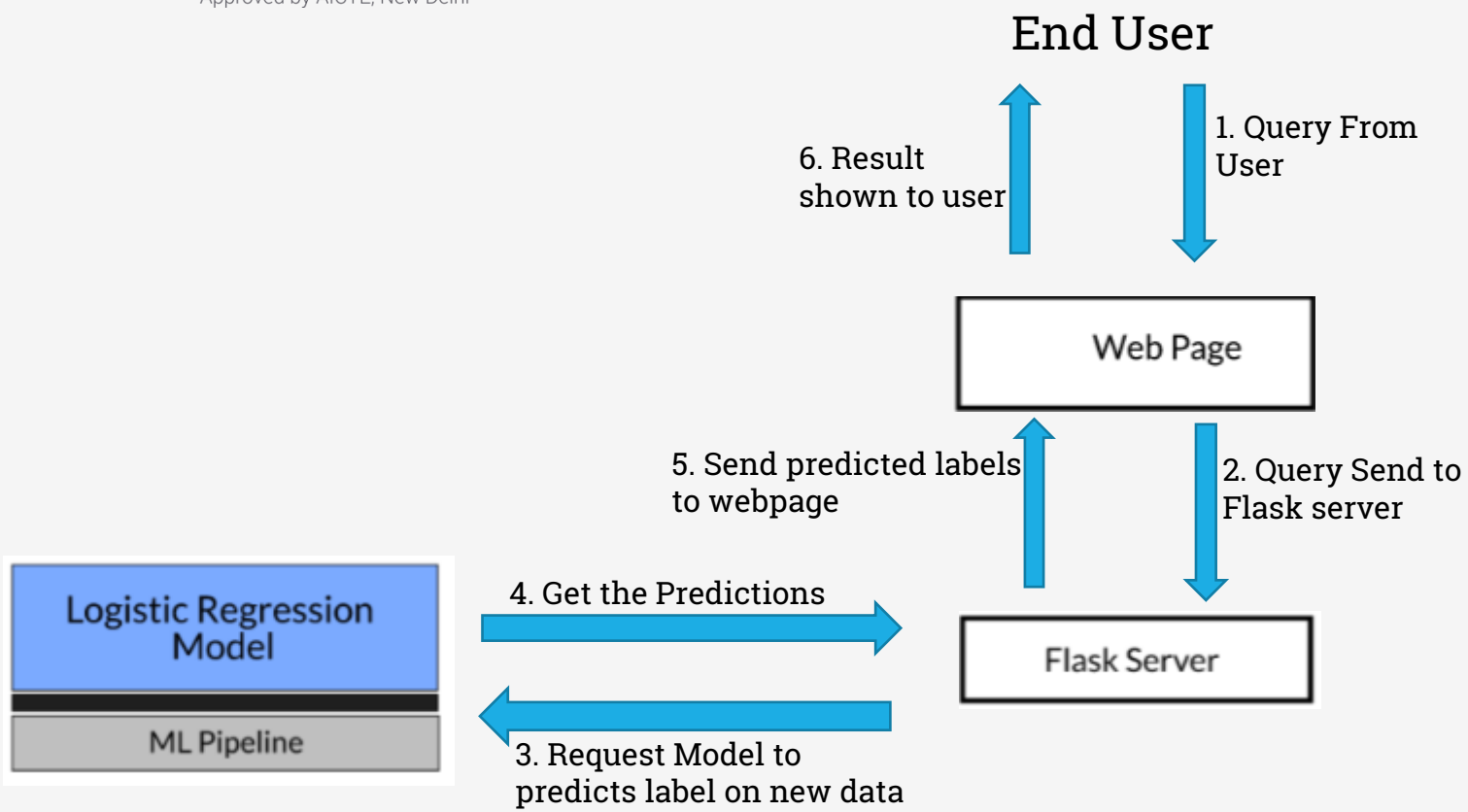
In all the models Logistic Regression is giving the best results , so we used the LR and divide the all the employees into three categories:

1. Low Risk (probabilities less than 0.3)
2. Medium Risk (probabilities between 0.3 to 0.7)
3. High Risk (probabilities between 1.0 to 0.7)



Model Deployment

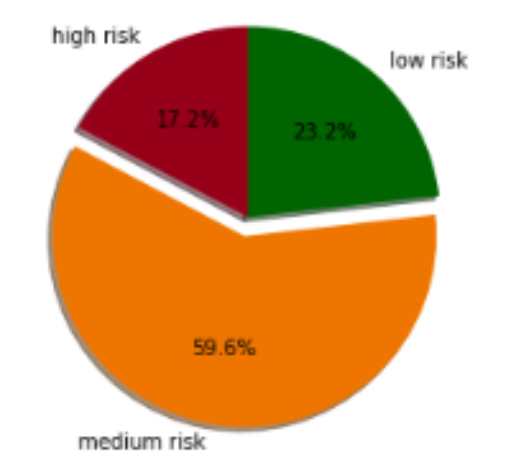
Demonstration



We are planning to deploy the model on the flask server with the help of pickel file of the saved model which can predict the category of Employee.

Results and Insights

Key Findings | Suggestions



It is found that most of the employees are on medium risk, around 60%, where as 23% are on lower risk, and only 17 % of employees are categorized as **high risk**.

The models have been developed and suggested some important steps to improve the compliance score of the employees.

1. With the help of this model, an organization can analyze the employee's previous data and predict if the employee is following the compliance or not.
2. For the improvement, an organization can take preventive steps to correct the behavior of the employee and avoid possible future losses.
3. Organizations can arrange important pieces of training for their employees to increase awareness of compliance in the business.

Conclusion and Future Work

Proposed solutions | Scope for future work

As we have limited dataset , we used some general Machine learning approach to asses the compliance risk of an employee.

In Future if we get sufficient data we can use the Advance Machine and Deep learning models to predict result with better accuracy and deploy the model on cloud based service e.g. AWS, GPC or Azure.

References

Bibliography | Webliography

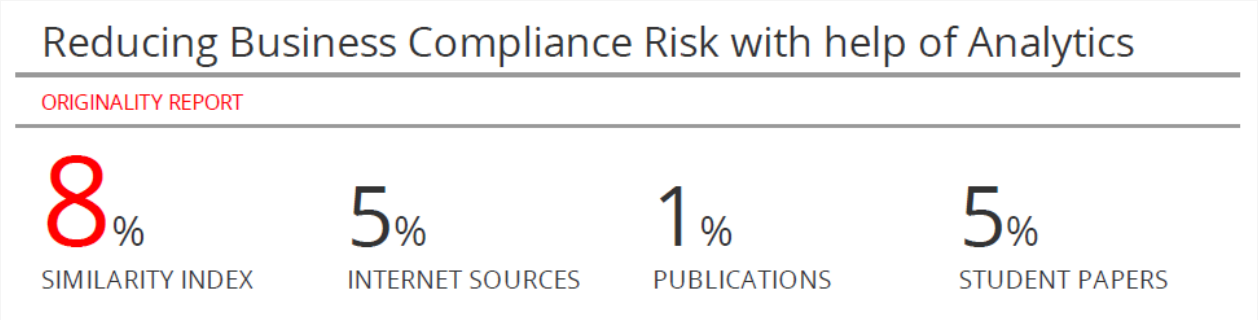
- Francesco. (2018). Sales Culture and Misconduct in the Financial Services Industry: An Analysis of Cross-selling Practices. Business Law Review. 39.
- Hennessy et, a. (2017). Federal Reserve Bank of New York Misconduct Risk, Culture, and Supervision.
- Mark Egan, G. M. (2019). The Market for Financial Adviser Misconduct. *University of Chicago*
- Namasivayam, K. (2006). Linking employee misbehavior to consumer satisfaction. *Journal of Foodservice Business Research* .
- Choi, J. H. (2019). Fraudulent Financial Reporting and the Consequences for Employees*. *SSRN Electronic Journal* .
- Sharma, S. (2018). The effect of training on employee performance. *International Journal of Recent Technology and Engineering* .
- Tracy, S. C. (2017). Federal Reserve Bank of New York Misconduct Risk, Culture, and Supervision. *www.gallup.com/poll/1597/Confidence-Institutions.aspx.*

Annexure

Additional Information | Plagiarism score

Similarity Index Report

- Software Used : **Turnitin**
- Date of Report Generation : **19- Aug-2022**
- Similarity Index in % : **8%**
- Total word count: **6,860**
- Name of the Guide : **Phaneendra Akula**



REVA
UNIVERSITY

Bengaluru, India

Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Thank You!