VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)
CHENNAI

Smart Internz

Vellore Institute of Technology, Chennai

Applied Data Science Externship

Project Title: - Risk Prediction in Corporate Financial Management Using IBM Auto AI Service

Project Report

-Submitted by

Team number: - 640

Ibrahim Ahmed Siddiqui 20BAI1189

Ratneshwar 20BAI1192

Table of Content

Introduction

Corporate financial management must include risk prediction since it enables businesses to anticipate possible risks and take smart countermeasures against them. Businesses can protect their financial stability, increase profitability, and guarantee long-term sustainability if they can effectively analyses risks. Companies now have the chance to use artificial intelligence (AI) and machine learning (ML) techniques to improve their capacity for risk prediction as a result of technological breakthroughs and the expanding availability of data.

Overview: -

The project titled "Risk Prediction in Corporate Financial Management Using IBM Auto AI Service" aims to develop a robust and efficient system for predicting financial risks in corporate settings. By utilizing the powerful capabilities of IBM Auto AI Service, this project seeks to leverage the vast amounts of financial and market data available to organizations, enabling them to proactively identify potential risks and take appropriate actions to mitigate them.

IBM Auto AI Service is a platform that combines the power of automated machine learning and AI to streamline the process of building, training, and deploying predictive models. It simplifies the complex task of developing predictive models by automating the feature engineering, algorithm selection, hyperparameter tuning, and model evaluation processes.

The project will involve several key steps.

- Collect the Data Set
- Create Waston Studio Project
- Add Auto AI Experiment
- Run Auto Ai Experiment
- Save the Model
- Deploy the Model

Once the models have been deployed, they can be utilized as independent tools or incorporated into current financial management systems to give real-time risk evaluations. Corporate decision-makers will be able to make data-driven decisions about investments, financial planning, and risk mitigation techniques thanks to these risk projections.

Purpose: -

The purpose of this project is to leverage the capabilities of IBM AutoAI to develop a system for predicting fraudulent transactions and create a web application that facilitates effective risk management in organizations. By automating the process of building predictive models and integrating them into a user-friendly interface, this project aims to achieve the following:

- Fraud Detection
- Risk Mitigation
- Enhanced Efficiency
- Financial Loss Reduction
- Improved Decision-Making
- Accurate Risk Assessment
- Timely Risk Identification
- Enhanced Financial Stability
- Improved Decision-Making
- Competitive Advantage

Literature Survey

Ma, X., Lv, S. Financial credit risk prediction in internet finance driven by machine learning. Neural Comput & Applic 31, 8359–8367 (2019). https://doi.org/10.1007/s00521-018-3963-6. This paper proposes a machine learning framework combining random forest and support vector machines to predict credit risk in corporate financial management. It demonstrates the effectiveness of the approach using real-world financial data.

"Deep learning for credit risk prediction: A comparison of deep neural networks and XGBoost" by Wang, H., Lu, Z., & Xu, G. (2019): The authors investigate the application of deep learning techniques, specifically deep neural networks, for credit risk prediction. They compare the performance of deep neural networks with XGBoost, a gradient boosting algorithm, using credit card default data.

Clintworth, M., Lyridis, D. & Boulougouris, E. Financial risk assessment in shipping: a holistic machine learning based methodology. Marit Econ Logist 25, 90–121 (2023). https://doi.org/10.1057/s41278-020-00183-2. The authors emphasize the importance of considering multiple risk factors in the shipping industry, such as market conditions, vessel performance, operational risks, and financial indicators. They propose a machine learning-based

approach that integrates these diverse factors to build predictive models for risk assessment.

Song, X.-P., Hu, Z.-H., Du, J.-G., & Sheng, Z.-H. (2014). Application of Machine Learning Methods to Risk Assessment of Financial Statement Fraud: Evidence from China. Journal of Forecasting, 33(7), 568-581. doi: 10.1002/for.2294. The authors investigate the effectiveness of machine learning techniques in identifying fraudulent financial statements. They utilize various machine learning algorithms to analyze a dataset of financial statement data from Chinese companies. The study aims to evaluate the predictive power of these algorithms in detecting fraudulent activities and to compare their performance with traditional statistical methods.

EXSISTING PROBLEM

Here are some financial risk management problems that might fail your strategy and drag you at the verge of facing potential financial risks:

- Economic factors.
- Your supplier's or external parties' decisions and actions.
- Financial instability in the market.
- Legal interventions.
- Internal call-to-actions.

PROPOSED SOLUTION

- Define all the high organizational level structures and product frames to support financial reporting.
- Simulate all the possible default cashflows and accruals for the financial risk management software.
- Quantify and manage the interest rate risk.

Theoretical Analysis

a. Data Preparation: The first step in risk prediction using IBM Auto AI service is to gather and prepare the relevant data. This includes financial data such as income statements, balance sheets, cash flow statements, and other relevant variables like market data, industry benchmarks, and macroeconomic indicators. The data should be cleaned, standardized, and transformed into a suitable format for analysis.

b. Feature Selection: Once the data is prepared, feature selection techniques can be applied to identify the most relevant variables for risk prediction. IBM Auto AI service employs automated feature selection methods, such as statistical tests, correlation analysis, and feature importance algorithms, to determine the subset of features that have the highest predictive power.

c. Model Building: IBM Auto AI service automates the process of building predictive models by selecting appropriate algorithms, tuning hyperparameters, and optimizing model performance. It explores a wide range of algorithms, including decision trees, random forests, gradient boosting, and neural networks, to find the best model for risk prediction.

d. Model Training and Validation: The selected models are trained using historical data and validated to ensure their accuracy and reliability. IBM Auto AI service utilizes techniques like cross-validation and hold-out validation to assess the performance of the models on unseen data. The models are evaluated based on metrics such as accuracy, precision, recall, and F1 score to gauge their effectiveness in risk prediction.

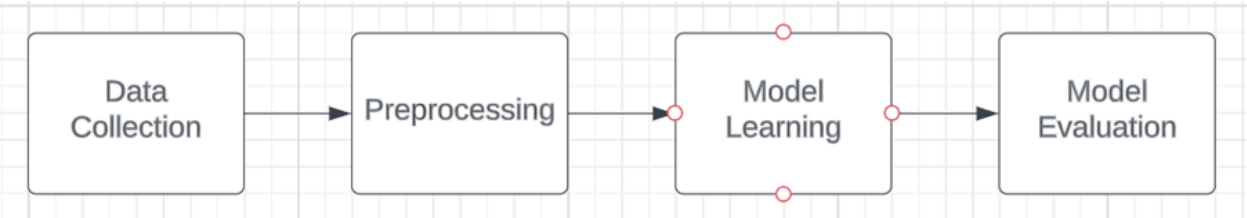
e. Risk Prediction and Visualization: Once the models are trained and validated, they can be used to predict future risks. IBM Auto AI service provides a user-friendly interface to input new data and generate risk predictions based on the trained models. The predictions can be visualized using charts, graphs, and other visual representations to facilitate better understanding and decision-making.

f. Risk Mitigation Strategies: The risk predictions generated by IBM Auto AI service can assist corporate financial managers in identifying potential risks and devising appropriate risk mitigation strategies. By understanding the factors contributing to the risks, managers can proactively implement measures to minimize their impact, such as adjusting financial strategies, diversifying

investments, implementing hedging techniques, or optimizing capital allocation.

g. Monitoring and Iteration: Risk prediction is an ongoing process, and it is crucial to continuously monitor the performance of the predictive models and refine them over time. IBM Auto AI service allows users to retrain and update the models as new data becomes available, ensuring that the risk predictions remain accurate and up to date.

Block Diagram: -



Hardware/Software Design: -

Minimum System Requirements: -

1. Processors: - Intel® Core™ i5 processor 4300M at 2.60 GHz or 2.59 GHz (1 socket, 2 cores, 2 threads per core), 8 GB of DRAMIntel® Xeon® processor ES-2698 v3 at 2.30 GHz (2 sockets, 16 cores each, I thread per core), 64 GB of DRAMIntel® Xeon PhiTM processor 7210 at 1.30 GHz (1 socket, 64 cores, 4 threads per core), 32 GB of DRAM, 16 GB of MCDRAM (flat mode enabled)
2. Disk space: 2 to 3 GB
3. Operating systems: Windows 7 or later, macOS, and Linux
4. Python versions: 3.9

Software: -

1. Web Browser
2. Vs Code
3. IBM Waston AI

4. Goggle Colab

Experimental Investigation

In the experimental investigation section of the project, we observe that,

Dataset Description: - The dataset captures various factors related to individuals' financial profiles, including age, gender, job, housing, savings, credit amount, duration, purpose, and risk classification. It provides a foundation for analyzing and predicting the risk involved in corporate financial management. The dataset consists of 1000 rows and 10 columns. Each row represents an individual and each column represents a specific attribute or feature of that individual. The goal is to predict the risk classification (good or bad) based on the given features.

Data Visualization: - We observe that,

- The count of good class was higher than that of bad class.
- People who own a house are more than being homeless or living in a rented house.
- Most of the people are skilled workers, with little savings account.
- Most of the people wanted loans of less than 5000 DM (Deutsche Mark).
- Many people want the loan for cars, radio, TV, furniture, and equipment, but few people want the loan for vacation, for purchasing domestic appliances or for repairs.
- Risk is lower for loan with purpose of purchase of radio/TV, but loans for vacation are mostly risky.
- Risk for males is comparatively lower than that compared to females.
- Smaller credit amount is generally asked for less duration, for 0 month to 50 months, while higher amount’s duration generally ranges from 35 months to 60 months.

Data Cleaning and Standardization: -

- We handle some missing value
- Remove Outliers
- Scaled the dataset

Baseline Models: - Implement and evaluate supervised algorithm to create a model. Algorithm like Logistics regression, Decision Tree, SVM, KNN, MLP

IBM Auto AI Model Building: - Utilize the IBM Auto AI service to build risk prediction models using the dataset. The service provides data preprocessing, feature engineering, model selection, and hyperparameter optimization and gives snap random forest model as a best model among many other models.

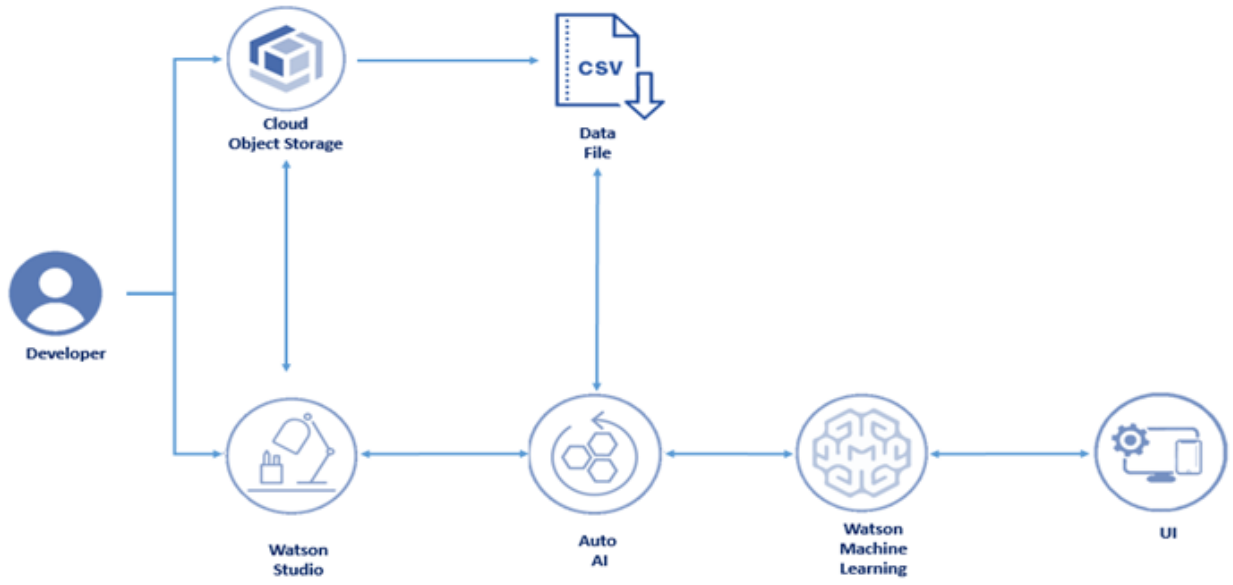
Model Evaluation: Evaluated the performance of the risk prediction models generated by IBM Auto AI based on accuracy and same with ml algorithms in ipynb file. Accuracy of different ml algo: -

a. Logistics Regression: - 74%
b. KNN: - 68%
c. Decision Tree: - 72%
d. SVM: - 73%
e. MLP: - 79%

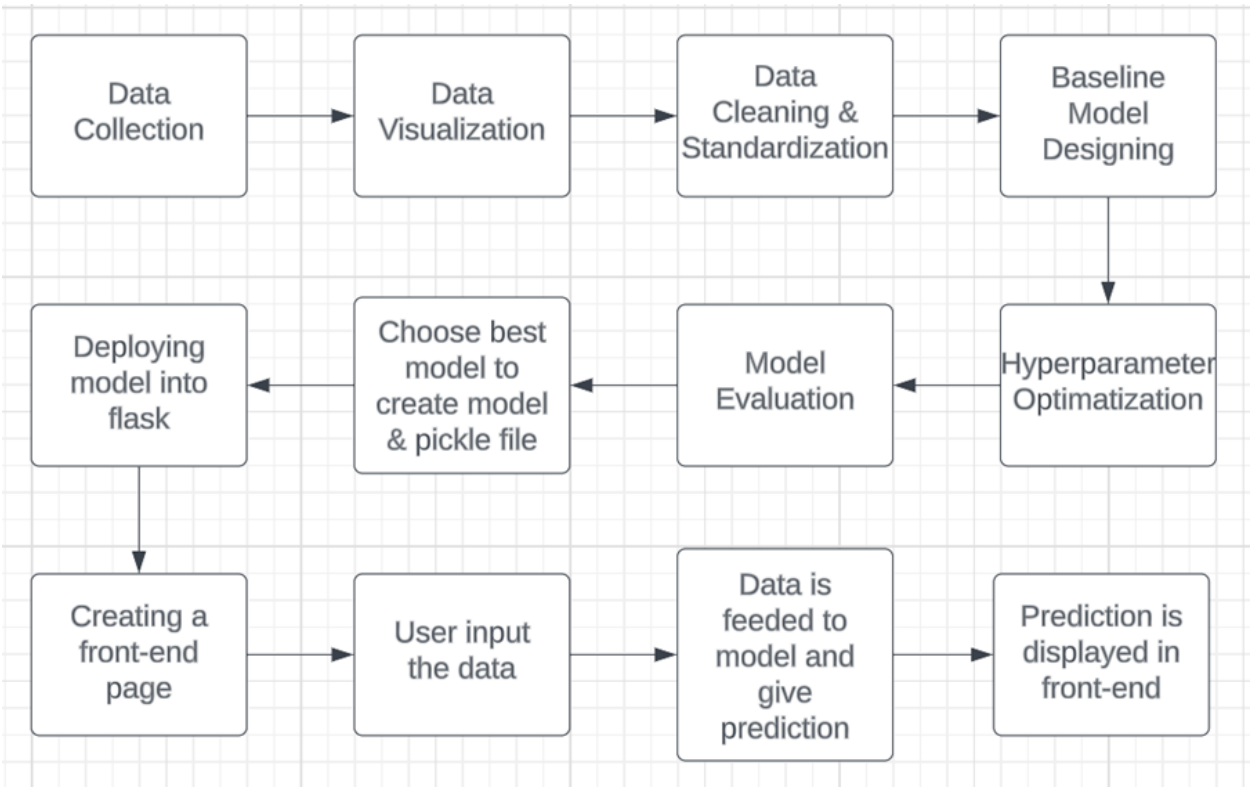
Interpretation of Results: MLP algorithm is given more accuracy than any other algorithm, so we will choose MLP as our model

Flow-Chart

For IBM Auto AI: -

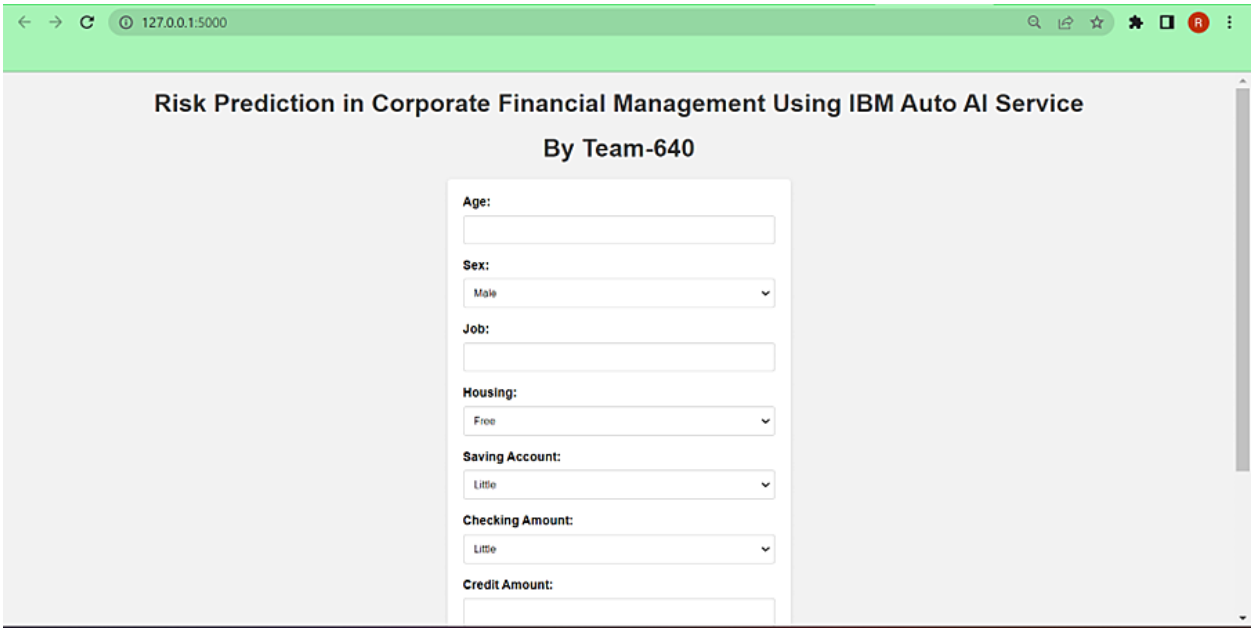


For Goggle Colab: -



Result

Website Screen shot: -



Form screen shot: -

Job:

Housing:

Free

Saving Account:

Little

Checking Amount:

Little

Credit Amount:

Duration:

Purpose:

Business

Predict

Filled form Screen shot: -

Age:

32

Sex:

Female

Job:

2

Housing:

Own

Saving Account:

Little

Checking Amount:

Moderate

Credit Amount:

2345

Duration:

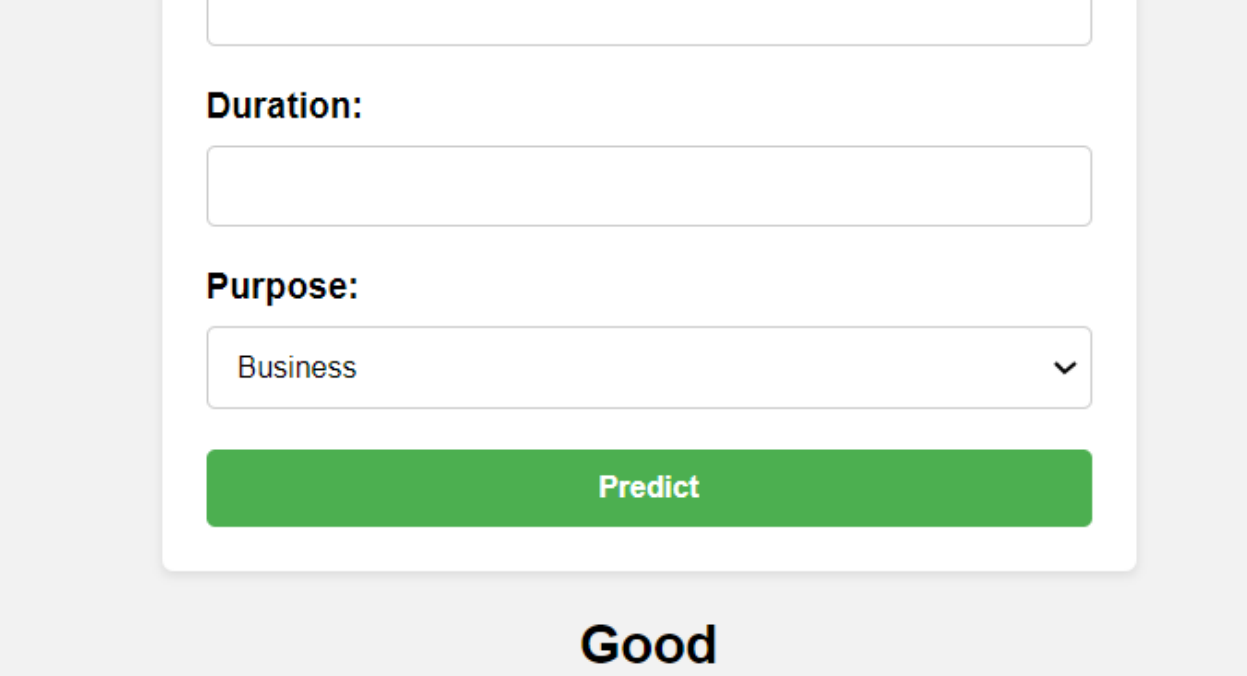
32

Purpose:

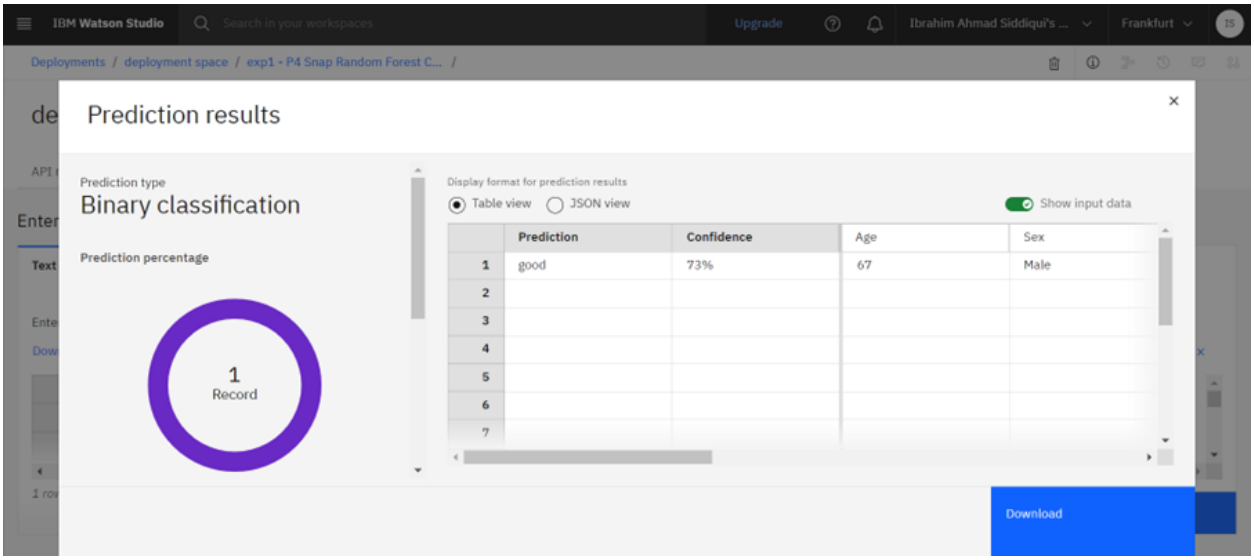
Car

Predict

Prediction Screen shot: -



Output from IBM Watson: -



Advantages & Disadvantages

Advantages of Risk Prediction in Corporate Financial Management Using IBM Auto AI: -

1. Enhanced risk assessment: The use of AI techniques, such as machine learning algorithms, can improve the accuracy and reliability of risk prediction in corporate financial management. It can identify potential risks more effectively and provide valuable insights for decision-making.

2. Efficient data analysis: AI-powered tools like IBM Auto AI can automate the process of data analysis, making it faster and more efficient. This can save time and resources for financial professionals, allowing them to focus on other critical tasks.

3. Scalability: AI models can handle large volumes of data and scale well to accommodate growing datasets. This scalability enables organizations to analyze complex financial data and make risk predictions on a larger scale.

4. Continuous learning and improvement: AI models can continuously learn and adapt based on new data. They can be updated and retrained periodically to improve their performance and capture changing risk patterns in the corporate financial landscape.

5. Objective decision-making: AI models provide an objective and data-driven approach to risk assessment, reducing potential biases that can arise from human judgment. This objectivity can lead to more consistent and reliable risk predictions.

Disadvantages of Risk Prediction in Corporate Financial Management Using IBM Auto AI:

1. Data quality and availability: The accuracy and effectiveness of AI models heavily rely on the quality and availability of data. If the input data is incomplete, inconsistent, or biased, it can affect the performance and reliability of the risk prediction model.

2. Interpretability and transparency: AI models, especially complex ones like deep learning algorithms, can lack interpretability. It may be challenging to understand and explain the underlying factors contributing to risk predictions, limiting the ability to gain insights and trust from stakeholders.

3. Overreliance on historical data: Risk prediction models are built based on historical data, which may not always capture future market dynamics or unexpected events. The models may struggle to adapt to novel risk scenarios or

sudden shifts in the financial landscape.

4. Cybersecurity and privacy risks: The use of AI in financial management involves handling sensitive and confidential data. It is crucial to implement robust cybersecurity measures to protect the data from unauthorized access or breaches, ensuring the privacy and integrity of the information.

5. Human expertise and judgment: While AI models can provide valuable risk predictions, they should not replace human expertise and judgment entirely. Financial professionals play a vital role in contextualizing the risk assessments, considering business-specific factors, and making informed decisions.

Application

1. Credit Risk Assessment: Financial institutions can use AI-based risk prediction models to assess the creditworthiness of borrowers. By analyzing historical data and various financial factors, the models can provide insights into the likelihood of default or delinquency, enabling more accurate credit risk evaluation.

2. Investment Risk Management: AI models can aid in investment risk management by analyzing market trends, historical data, and economic indicators. Financial professionals can leverage these predictions to make informed investment decisions, optimize portfolios, and mitigate potential risks.

3. Fraud Detection: AI-based risk prediction models can help identify potential fraudulent activities in corporate financial transactions. By analyzing patterns and anomalies in financial data, the models can flag suspicious transactions or activities, enhancing fraud detection and prevention measures.

4. Financial Forecasting: Risk prediction models can assist in financial forecasting by analyzing historical financial data and market trends. Organizations can

leverage these predictions to anticipate potential risks and uncertainties, enabling better financial planning and decision-making.

5. Regulatory Compliance: AI models can support organizations in complying with regulatory requirements in the financial sector. By analyzing data and identifying potential compliance risks, organizations can take proactive measures to mitigate these risks and ensure adherence to regulatory guidelines.

6. Merger and Acquisition Analysis: Risk prediction models can be utilized in assessing the risks associated with mergers and acquisitions. By analyzing financial data and market trends, the models can help identify potential risks and synergies, aiding in the decision-making process.

7. Risk-Based Pricing: Financial institutions can use risk prediction models to determine risk-based pricing strategies. By accurately assessing the risks associated with lending or insurance products, organizations can offer personalized pricing options that align with the risk profiles of individual customers.

8. Portfolio Risk Management: AI models can assist in portfolio risk management by analyzing the risk profiles of individual assets and optimizing portfolio allocations. This helps financial institutions optimize risk-return trade-offs and achieve diversification objectives.

9. Stress Testing: Risk prediction models can be used in stress testing scenarios to evaluate the resilience of financial systems or portfolios under adverse market conditions. By simulating various risk scenarios, organizations can assess their vulnerability and take preventive measures.

10. Risk Reporting and Monitoring: AI-based risk prediction systems can provide real-

time risk reporting and monitoring capabilities. This enables organizations to track risk exposures, identify emerging risks, and take timely actions to mitigate them

Conclusion

In conclusion, the project "Risk Prediction in Corporate Financial Management Using IBM Auto AI" aimed to develop a predictive model using IBM Auto AI and Goggle Colab Notebook to assess the risk in corporate financial management. The project involved several steps including data preprocessing, feature selection, model training, and evaluation.

The dataset used for the project consisted of various features such as age, sex, job, housing, saving accounts, checking account, credit amount, duration, purpose, and risk. The dataset was analyzed and processed to handle missing values and encode categorical variables.

The best-performing model was trained using the MLP algorithm. The model achieved an accuracy of 79% in predicting the risk associated with corporate financial management. The precision and recall scores were also evaluated, with class 0 (low risk) having a precision of 0.74 and recall of 0.38, and class 1 (high risk) having a precision of 0.80 and recall of 0.95. These scores indicate that the model performs relatively well in identifying high-risk cases, while there is room for improvement in correctly identifying low-risk cases.

Overall, the project successfully demonstrated the use of IBM Auto AI in predicting the risk in corporate financial management. The results can be utilized by financial professionals and organizations to make informed decisions and take appropriate measures to mitigate potential risks. However, further improvements can be made by exploring other algorithms, optimizing model parameters, and considering additional features or data sources to enhance the predictive performance.

Future Scope

As technology advances and more data becomes available, the accuracy and capabilities of AI-driven risk prediction models are expected to improve. The integration of advanced techniques such as natural language processing, deep learning, and reinforcement learning can enhance risk assessment and prediction accuracy. It presents several opportunities for future enhancements and advancements. Here are some potential areas for future scope:

a. Incorporating additional features: The existing dataset used in the project included a limited set of features. Future work can involve exploring and incorporating additional relevant features that could contribute to more accurate

risk prediction. This could include financial ratios, industry-specific indicators, market trends, or macroeconomic factors.

b. Fine-tuning model parameters: Although the project achieved satisfactory results with the selected model parameters, there is room for further optimization. Conducting a comprehensive parameter tuning process using techniques like grid search or Bayesian optimization can help identify the best combination of parameters for improved model performance.

c. Ensemble modeling: Ensemble techniques, such as combining multiple models or implementing stacking/blending approaches, can be explored to enhance the predictive power of the risk prediction model. By leveraging the strengths of different algorithms or models, ensemble methods have the potential to provide more robust and accurate risk assessments.

d. Real-time risk monitoring: Building a real-time risk monitoring system can be a valuable future extension. By continuously collecting and analyzing financial data, the system can provide dynamic risk assessments and alerts in corporate financial management. This can aid in proactive decision-making and timely risk mitigation.

e. Integration with external data sources: Augmenting the existing dataset with external data sources, such as financial market data, industry reports, or news sentiment analysis, can provide a broader context for risk prediction. Incorporating such external data can enhance the accuracy and relevance of risk assessments.

f. Interpretability and explainability: Developing techniques to interpret and explain the risk prediction model's decisions can increase stakeholders' trust and understanding. Employing methods like feature importance analysis, model explainability algorithms, or generating decision rules can help provide transparent and interpretable risk assessments.

g. Continuous model evaluation and updates: As new data becomes available; it is essential to continuously evaluate and update the risk prediction model. Monitoring the model's performance over time, retraining with fresh data, and validating against actual risk outcomes can help ensure its effectiveness and reliability.

By focusing on these future directions, the project can advance the field of risk prediction in corporate financial management, providing more accurate, timely, and actionable insights for businesses and financial professionals.

Bibliography

a. Journal Article:

i. Ma, X., Lv, S. Financial credit risk prediction in internet finance driven by machine learning. Neural Comput & Applic 31, 8359–8367 (2019). https://doi.org/10.1007/s00521-018-3963-6

ii. "Deep learning for credit risk prediction: A comparison of deep neural networks and XGBoost" by Wang, H., Lu, Z., & Xu, G. (2019):

iii. Clintworth, M., Lyridis, D. & Boulougouris, E. Financial risk assessment in shipping: a holistic machine learning based methodology. Marit Econ Logist 25, 90–121 (2023). https://doi.org/10.1057/s41278-020-00183-2

iv. Song, X.-P., Hu, Z.-H., Du, J.-G., & Sheng, Z.-H. (2014). Application of Machine Learning Methods to Risk Assessment of Financial Statement Fraud: Evidence from China. Journal of Forecasting, 33(7), 568-581. doi: 10.1002/for.2294

b. Website:

i. IBM Cloud

ii. Greek for Greek

iii. Google Colab

iv.