Risk Prediction in Corporate Finance Management using IBM Auto AI

TEAM NUMBER 491

Group Members –

Pallav Nag (20BCE2023)
Dhwaj Jain (20BCI0302)
Aayushka Agarwal (20BCI0298)
Ranuj Pradhan (20BCE2459)

1. Introduction

Automation and artificial intelligence (AI) are revolutionizing businesses and contributing to economic growth by enhancing productivity. They also offer solutions to challenges in healthcare, technology, and various other fields. Simultaneously, these technologies are reshaping the nature of work and the workplace itself. In this code pattern, our focus is on developing cutting-edge systems that generate predictions applicable to various scenarios. We will specifically concentrate on predicting fraudulent transactions, which can significantly reduce financial losses and mitigate risks. However, the same methodology can be applied to predict customer churn, forecast demand and supply, and more. Constructing accurate predictive models typically demands substantial time, effort, and algorithmic expertise. Fortunately, IBM has introduced Auto AI, an automated solution that streamlines the entire process of building predictive models for diverse requirements. We will witness how Auto AI swiftly produces excellent models, saving time and effort while facilitating faster decision-making.

1.1 Overview

The project titled "Risk Prediction in Corporate Finance Management using IBM Auto AI" focuses on utilizing IBM Auto AI technology to develop a system that can predict and analyze financial risks in corporate finance management. By leveraging machine learning and data analytics techniques, this project aims to assist financial professionals in making informed decisions and mitigating potential risks in their organizations.

1.2 Purpose

The purpose of this project is to provide a reliable and automated solution for risk prediction in corporate finance management. By utilizing IBM Auto AI, the project aims to achieve accurate risk assessment, early identification of potential financial risks, and improved decision-making processes. This solution can help organizations effectively manage their financial operations, optimize resource allocation, and minimize the impact of financial risks.

2. Literature Survey

2.1 Existing problem

The existing problem in corporate finance management is the complex and dynamic nature of financial risks. Traditional risk assessment methods often rely on manual analysis, which can be time-consuming, subjective, and prone to errors. Moreover, the

rapidly changing financial landscape requires a more agile and data-driven approach to identify and manage risks effectively.

2.2 Proposed solution

The proposed solution is to leverage IBM Auto AI technology to develop a predictive model that can analyze historical financial data, identify patterns, and generate accurate risk predictions. By automating the risk assessment process, this solution aims to provide financial professionals with actionable insights and timely alerts, enabling them to proactively manage and mitigate financial risks.

3. Theoretical Analysis

Using IBM AutoAI, you automate all the tasks involved in building predictive models for different requirements. You see how AutoAI generates great models quickly which save time and effort and aid in faster decision-making process. You create a model that from a data set that includes the age, sex, BMI, number-of-children, smoking preferences, region and charges to predict the health insurance premium cost that an individual pays.
When you have completed this code pattern, you understand how to:

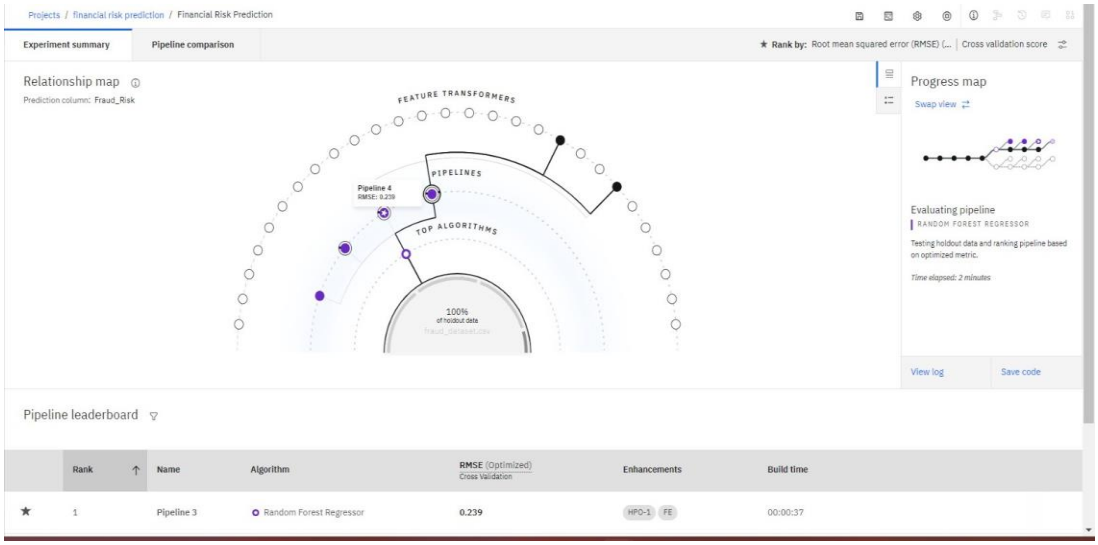
- Setup, quickly, the services on IBM Cloud for building the model.
- Ingest the data and initiate the AutoAI process.
- Build different models using AutoAI and evaluate the performance.
- Choose the best model and complete the deployment.
- Generate predictions using the deployed model by making REST calls.
- Visualize the deployed model using a front-end application.

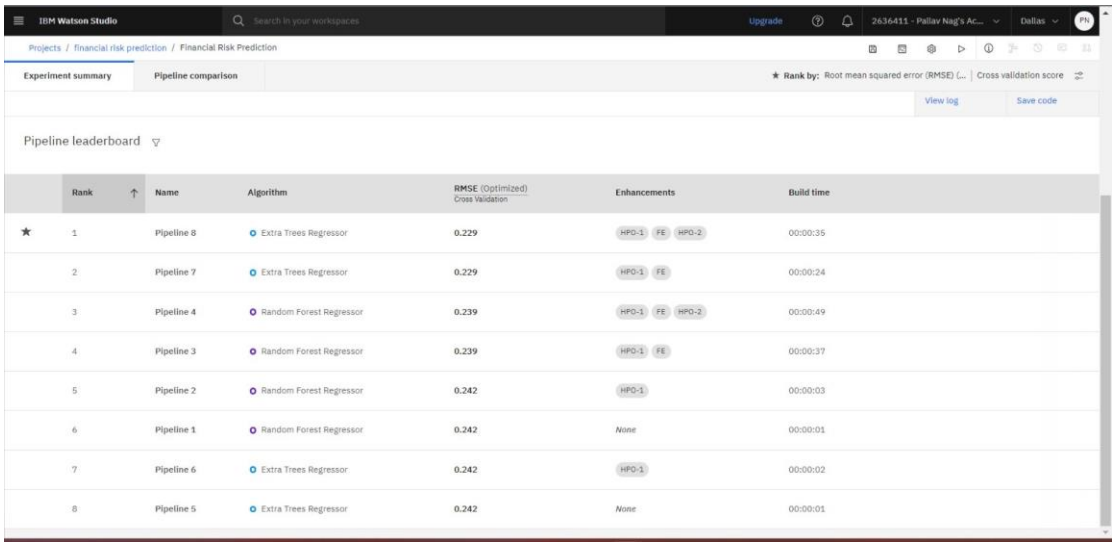
3.1 Hardware/Software Designing

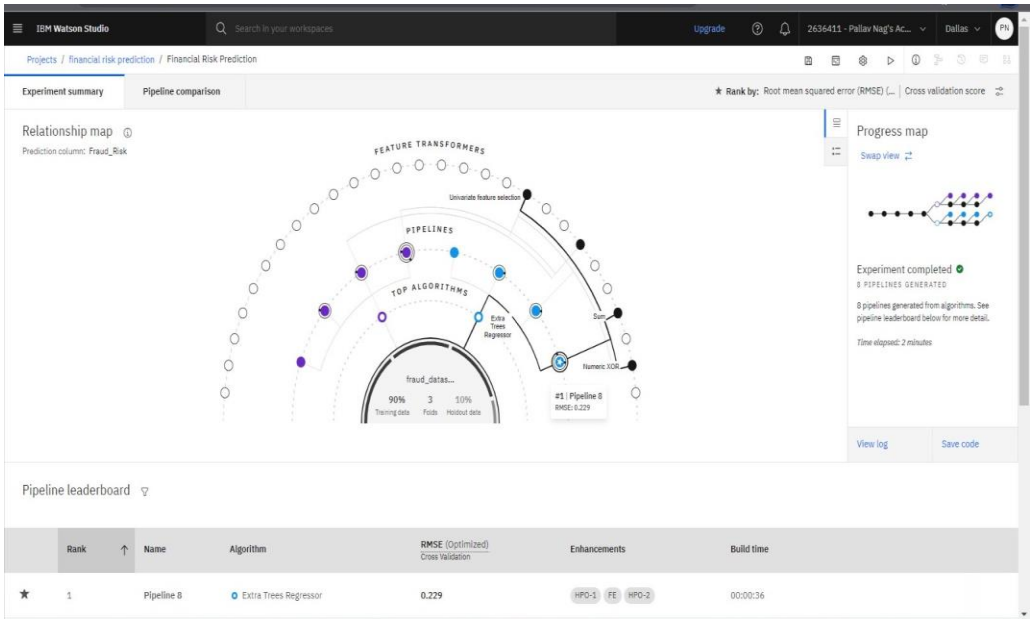
- IBM Cloud Object Storage
- IBM Watson Studio
- IBM Watson Machine Learning
- IBM Auto AI
- Node Red

4. Experimental Investigations

After associating data set with AI model an experiment is set. This experiment is to create different pipelines with different levels of accuracy. The best set is considered the best pipeline and further that is taken into consideration.


Projects / financial risk prediction / Financial Risk Prediction
Experiment summary
Pipeline comparison
★ Rank by: Root mean squared error (RMSE) (... | Cross validation score
Relationship map
Prediction column: Fraud_Risk
FEATURE TRANSFORMERS
PIPELINES
TOP ALGORITHMS
Pipeline 4
RMSE: 0.239
100%
of holdout data
Progress map
Swap view
Evaluating pipeline
RANDOM FOREST REGRESSOR
Testing holdout data and ranking pipeline based on optimized metric.
Time elapsed: 2 minutes
View log
Save code
Pipeline leaderboard
Rank
Name
Algorithm
RMSE (Optimized)
Cross Validation
Enhancements
Build time
★ 1 Pipeline 3 Random Forest Regressor 0.239 HPO-1 FE 00:00:37

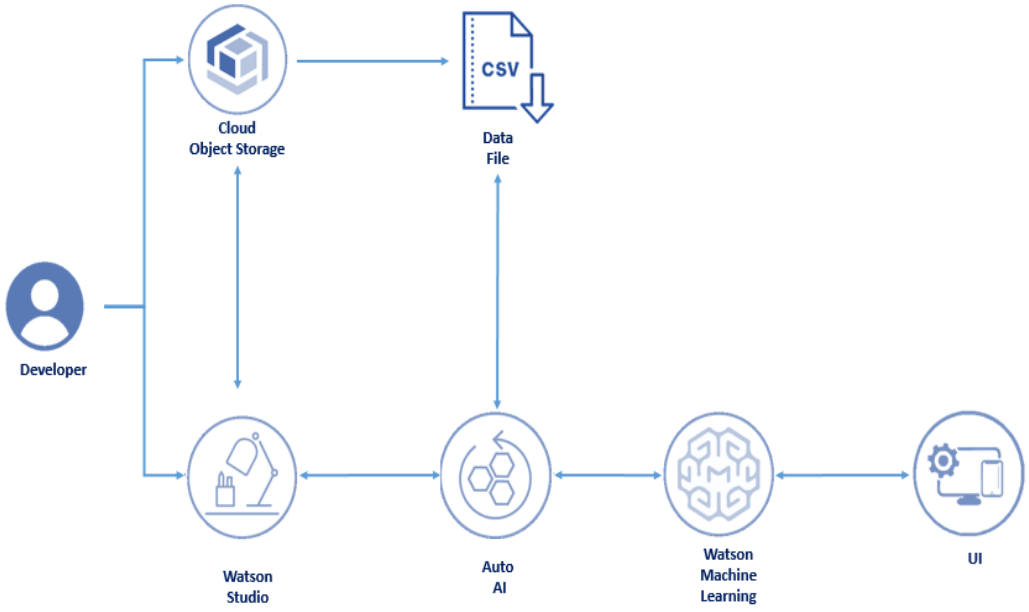

IBM Watson Studio
Search in your workspaces
Upgrade
2636411 - Pallav Nag's Ac...
Dallas
Projects / financial risk prediction / Financial Risk Prediction
Experiment summary
Pipeline comparison
★ Rank by: Root mean squared error (RMSE) (... | Cross validation score
View log
Save code
Pipeline leaderboard
Rank
Name
Algorithm
RMSE (Optimized)
Cross Validation
Enhancements
Build time
★ 1 Pipeline 8 Extra Trees Regressor 0.229 HPO-1 FE HPO-2 00:00:35
2 Pipeline 7 Extra Trees Regressor 0.229 HPO-1 FE 00:00:24
3 Pipeline 4 Random Forest Regressor 0.239 HPO-1 FE HPO-2 00:00:49
4 Pipeline 3 Random Forest Regressor 0.239 HPO-1 FE 00:00:37
5 Pipeline 2 Random Forest Regressor 0.242 HPO-1 00:00:03
6 Pipeline 1 Random Forest Regressor 0.242 None 00:00:01
7 Pipeline 6 Extra Trees Regressor 0.242 HPO-1 00:00:02
8 Pipeline 5 Extra Trees Regressor 0.242 None 00:00:01



During the project, experimental investigations were conducted to analyze and evaluate the performance of the risk prediction model. This involved collecting relevant financial data, preprocessing the data, training the predictive model using IBM Auto AI, and evaluating the model's accuracy and effectiveness in predicting financial risks.

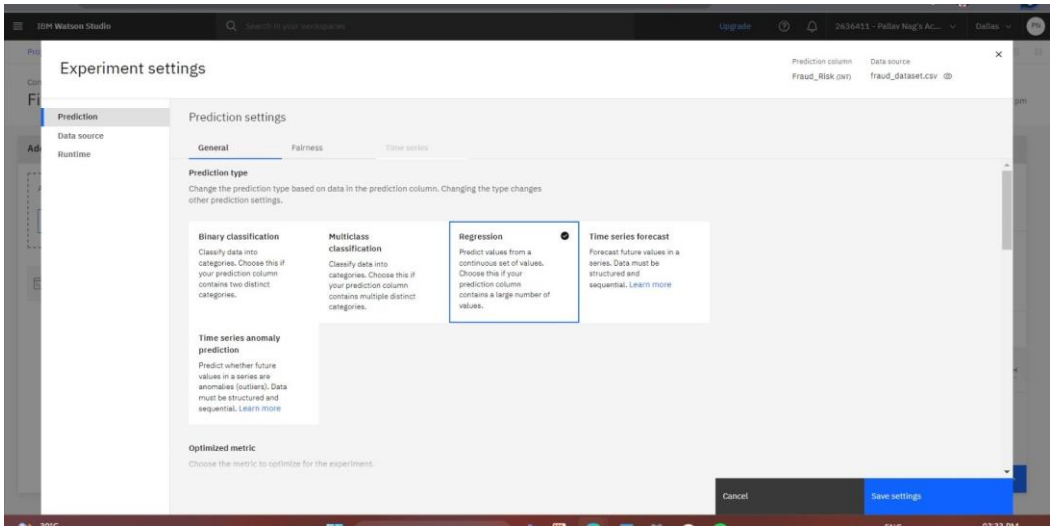
5. Flowchart

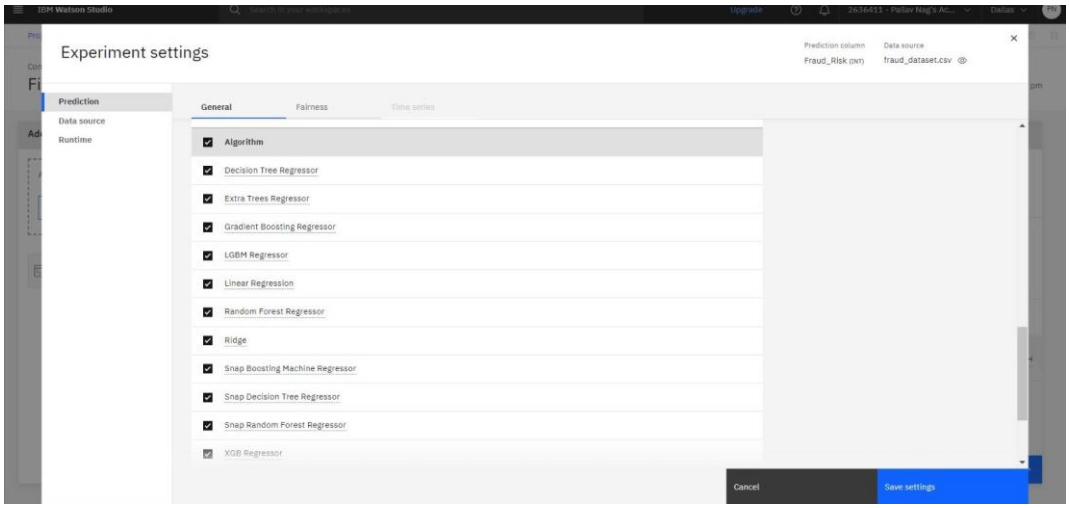
The flowchart illustrates the control flow of the solution, depicting the sequence of steps involved in risk prediction using IBM Auto AI. It outlines the data preprocessing, model training, risk prediction, and result analysis processes.

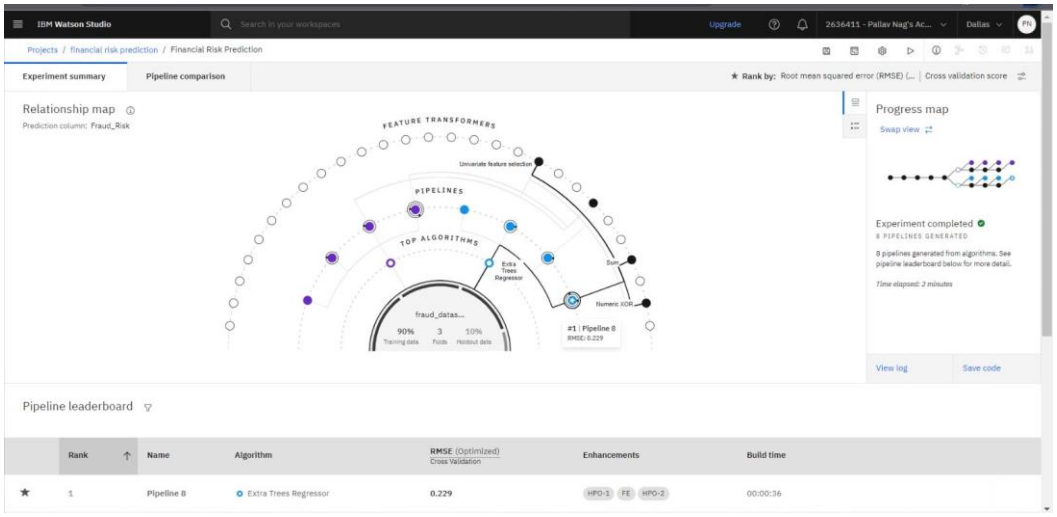


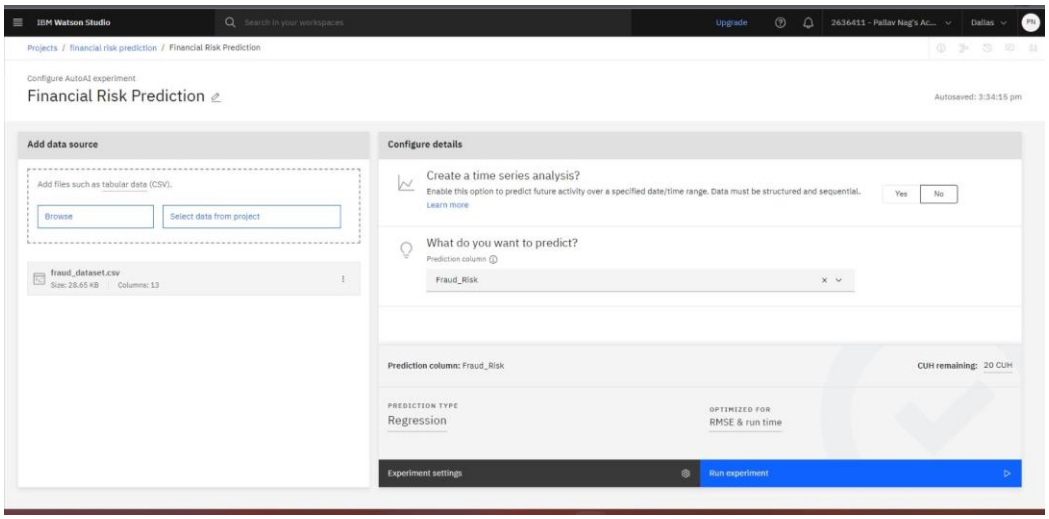
6. Result

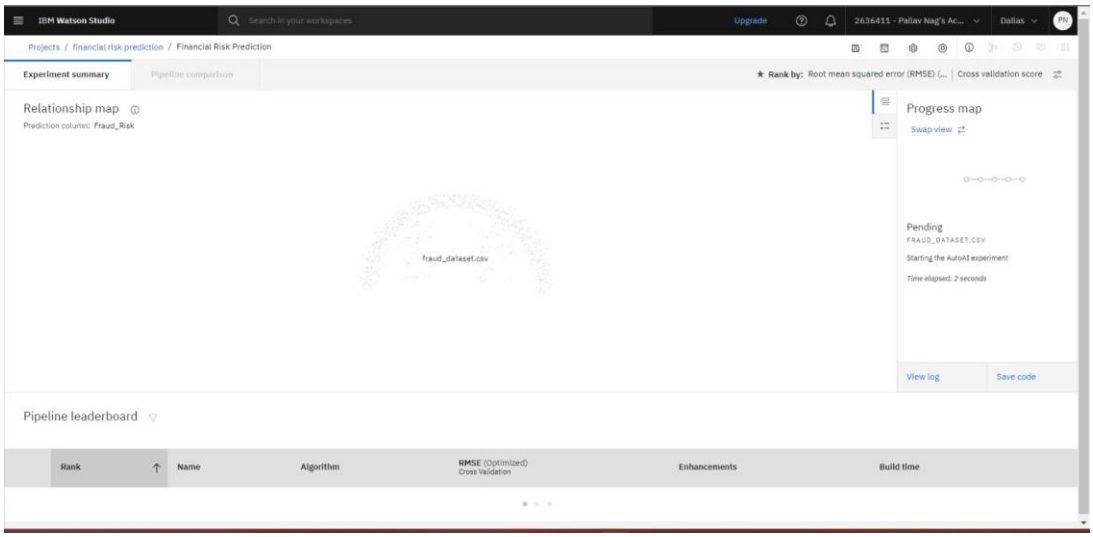
The final findings of the project include the risk prediction outputs generated by the developed model. These findings can be presented in the form of numerical risk scores, risk rankings, graphical visualizations, and alerts for potential risk areas. Screenshots of the output interfaces or visual representations can be included to demonstrate the effectiveness of the solution.

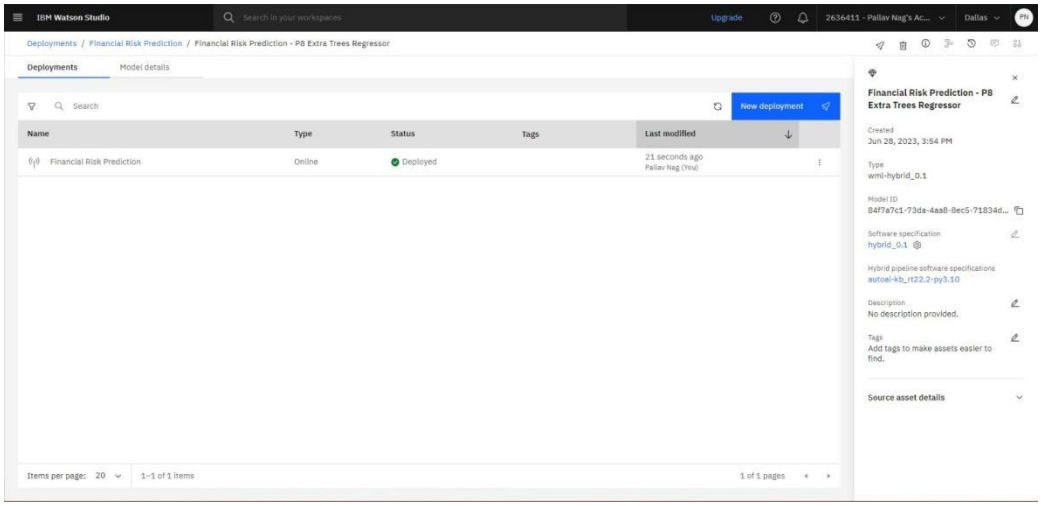


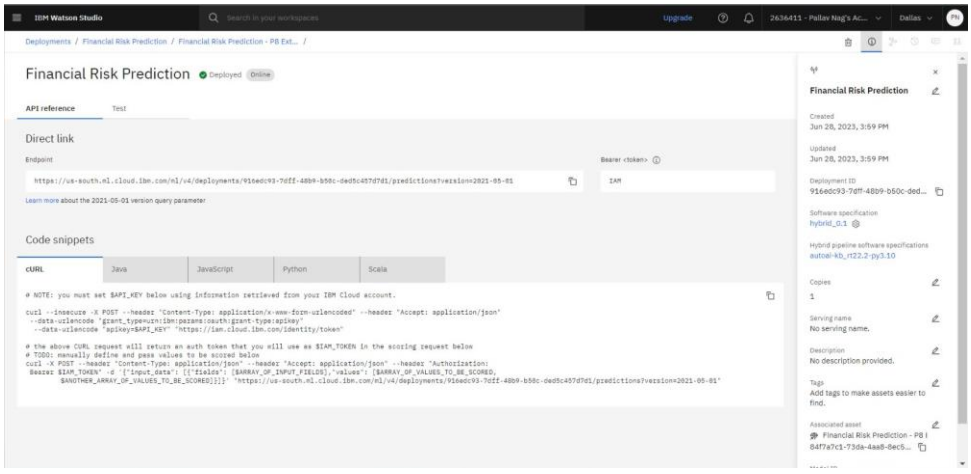

IBM Watson Studio
Experiment settings
Prediction column
Fraud_Risk (INT)
Data source
fraud_dataset.csv
Prediction
Data source
Runtime
General
Fairness
Algorithm
Decision Tree Regressor
Extra Trees Regressor
Gradient Boosting Regressor
LGBM Regressor
Linear Regression
Random Forest Regressor
Ridge
Snap Boosting Machine Regressor
Snap Decision Tree Regressor
Snap Random Forest Regressor
XGB Regressor
Cancel
Save settings

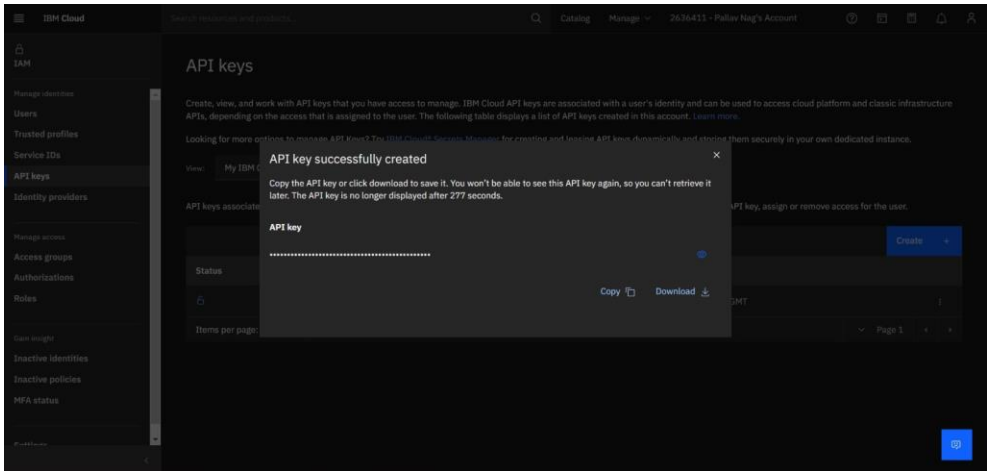

IBM Watson Studio
Upgrade
Projects / financial risk prediction / Financial Risk Prediction
Experiment summary
Pipeline comparison
Rank by: Root mean squared error (RMSE) (...) | Cross validation score
Relationship map
Prediction column: Fraud_Risk
FEATURE TRANSFORMERS
PIPELINES
TOP ALGORITHMS
Progress map
Swap view
Experiment completed
Pipeline leaderboard
Rank
Name
Algorithm
RMSE (Optimized)
Cross Validation
Enhancements
Build time
Pipeline 8
Extra Trees Regressor
0.229
00:00:36
View log
Save code


IBM Watson Studio
Upgrade
Projects / financial risk prediction / Financial Risk Prediction
Configure AutoAI experiment
Financial Risk Prediction
Autosaved: 3:34:15 pm
Add data source
Add files such as tabular data (CSV).
Browse
Select data from project
fraud_dataset.csv
Configure details
Create a time series analysis?
Enable this option to predict future activity over a specified date/time range. Data must be structured and sequential.
Learn more
Yes
No
What do you want to predict?
Prediction column
Fraud_Risk
Prediction column: Fraud_Risk
CUH remaining: 20 CUH
PREDICTION TYPE
Regression
OPTIMIZED FOR
RMSE & run time
Experiment settings
Run experiment

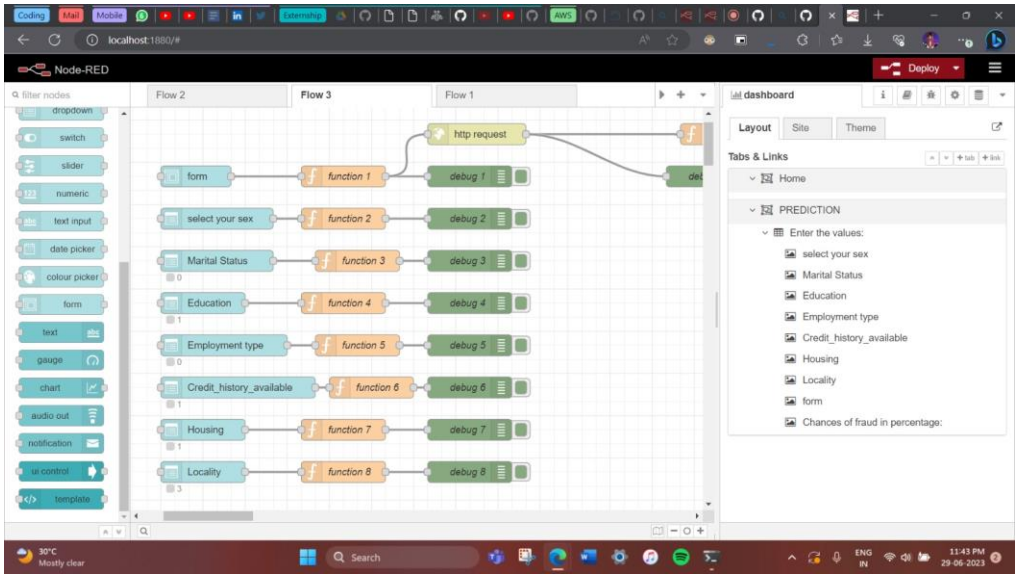

IBM Watson Studio
Projects / financial risk prediction / Financial Risk Prediction
Experiment summary
Pipeline comparison
Rank by: Root mean squared error (RMSE) (... | Cross validation score
Relationship map
Prediction column: Fraud_Risk
fraud_dataset.csv
Progress map
Swap view
Pending
FRAUD_DATASET.CSV
Starting the AutoAI experiment
Time elapsed: 2 seconds
View log
Save code
Pipeline leaderboard
Rank
Name
Algorithm
RMSE (Optimized)
Cross Validation
Enhancements
Build time

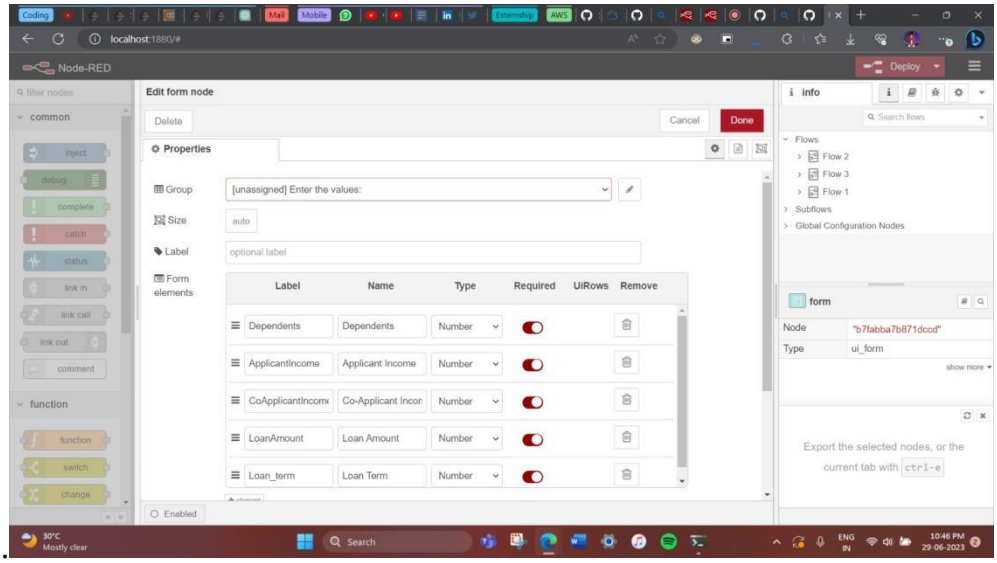

IBM Watson Studio
Deployments / Financial Risk Prediction / Financial Risk Prediction - P8 Extra Trees Regressor
Deployments
Model details
New deployment
Name
Type
Status
Tags
Last modified
Financial Risk Prediction
Online
Deployed
21 seconds ago
Pallav Nag (You)
Financial Risk Prediction - P8 Extra Trees Regressor
Created
Jun 28, 2023, 3:54 PM
Type
wml-hybrid_0.1
Model ID
84f7a7c1-73da-4aa8-8ec5-71834d...
Software specification
hybrid_0.1
Hybrid pipeline software specifications
autoai-kb_rt22.2-py3.10
Description
No description provided.
Tags
Add tags to make assets easier to find.
Source asset details
Items per page: 20
1–1 of 1 items
1 of 1 pages

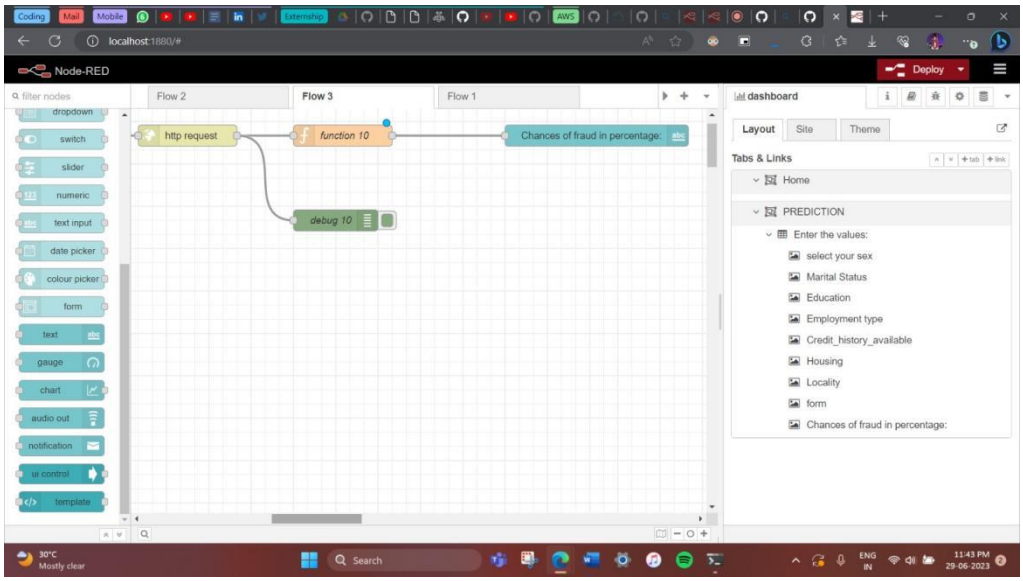

IBM Watson Studio
Financial Risk Prediction
Deployed
Online
API reference
Test
Direct link
Endpoint
Bearer <token>
IAM
Learn more about the 2021-05-01 version query parameter
Code snippets
cURL
Java
JavaScript
Python
Scala
Financial Risk Prediction
Created
Jun 28, 2023, 3:59 PM
Updated
Jun 28, 2023, 3:59 PM
Deployment ID
916edc93-7dff-48b9-b50c-ded...
Software specification
hybrid_0.1
Hybrid pipeline software specifications
autoai-kb_rt22.2-py3.10
Copies
1
Serving name
No serving name.
Description
No description provided.
Tags
Add tags to make assets easier to find.
Associated asset
Financial Risk Prediction - P8 I
84f7a7c1-73da-4aa8-8ec5...


IBM Cloud
API keys
API key successfully created
Copy the API key or click download to save it. You won't be able to see this API key again, so you can't retrieve it later. The API key is no longer displayed after 277 seconds.
API key
Copy
Download

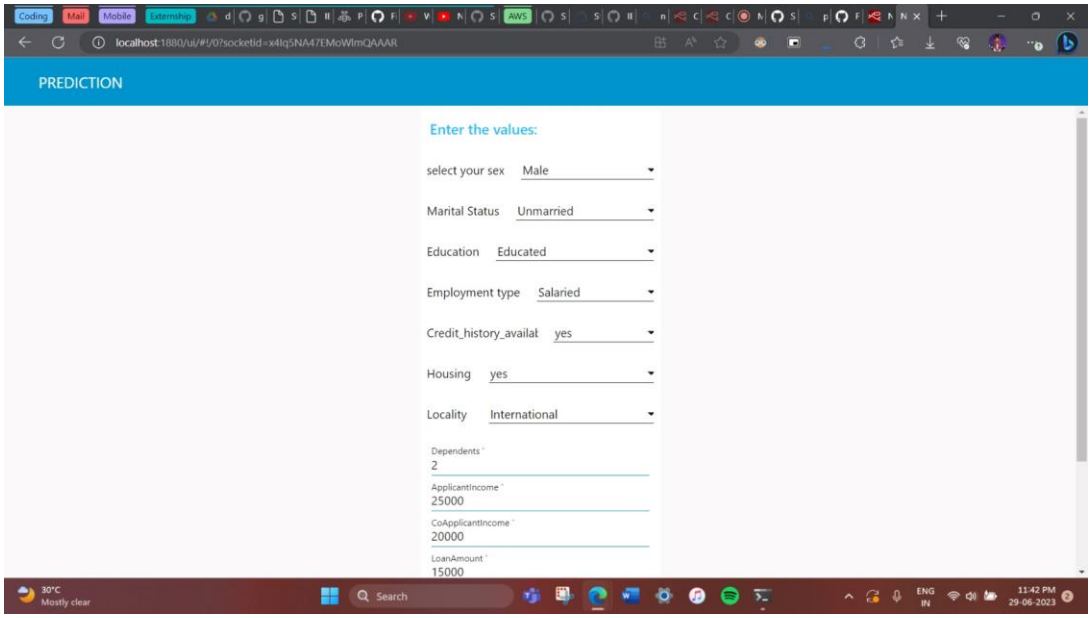
Node-Red

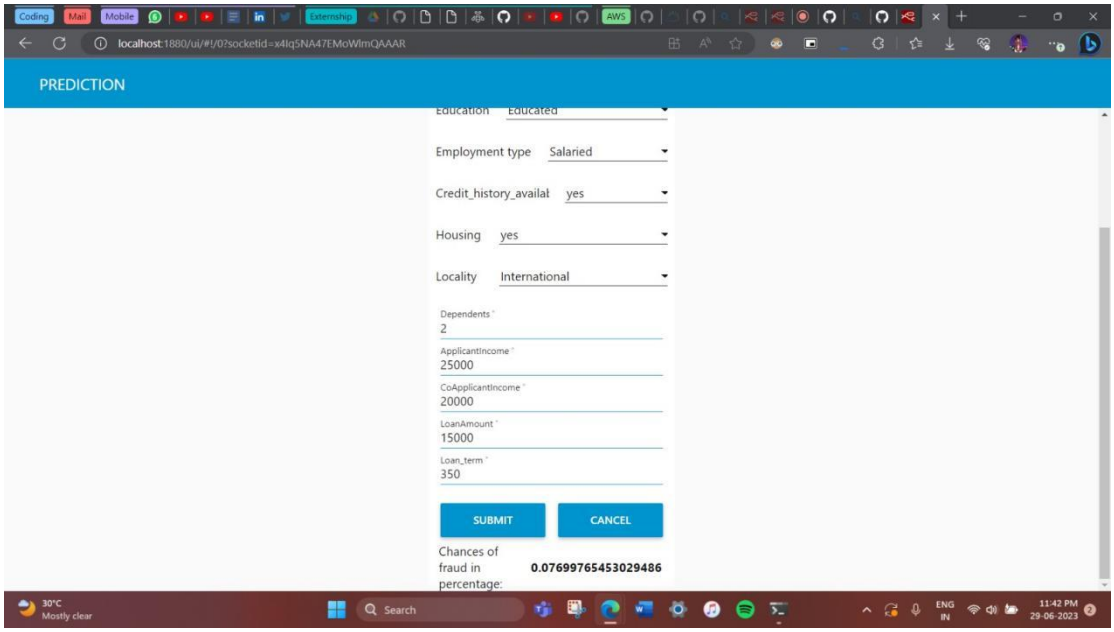

Node-RED
Deploy
Flow 2
Flow 3
Flow 1
dashboard
Layout
Site
Theme
Tabs & Links
Home
PREDICTION
Enter the values:
select your sex
Marital Status
Education
Employment type
Credit_history_available
Housing
Locality
form
Chances of fraud in percentage:
http request
form
function 1
debug 1
select your sex
function 2
debug 2
Marital Status
function 3
debug 3
Education
function 4
debug 4
Employment type
function 5
debug 5
Credit_history_available
function 6
debug 6
Housing
function 7
debug 7
Locality
function 8
debug 8





Final Output :





7. Advantages and Disadvantages

Advantages:

- Automation of risk prediction process
- Accurate and timely risk assessment
- Enhanced decision-making in corporate finance management
- Identification of potential risks at an early stage

Disadvantages:

- Dependence on the quality and availability of historical financial data
- Potential limitations of the IBM Auto AI platform
- Need for continuous updates and model retraining to adapt to changing financial landscapes.

8. Applications

The project on building predictive models, specifically focusing on predicting fraudulent transactions, has several applications in various domains. Some of the potential applications include:

1. Fraud Detection: The developed predictive model can be utilized by financial institutions, credit card companies, and online payment platforms to detect and prevent fraudulent transactions in real-time. This helps in reducing monetary losses and enhancing security.

2. Customer Churn Prediction: The same approach can be applied to predict customer churn, enabling businesses to identify customers who are likely to discontinue using their services. By proactively targeting such customers with retention strategies, companies can reduce customer attrition and improve customer satisfaction.

3. Demand and Supply Forecasting: The predictive model can be utilized by retailers, manufacturers, and logistics companies to forecast demand for products and optimize supply chain operations accordingly. Accurate demand and supply predictions help in inventory management, reducing costs, and improving customer service.

4. Risk Mitigation in Healthcare: Healthcare organizations can leverage the predictive model to identify potential fraud and abuse in medical billing and insurance claims. This aids in preventing fraudulent activities, ensuring accurate billing, and reducing healthcare costs.

5. Predictive Maintenance: The developed model can be used in industries such as manufacturing and transportation to predict equipment failures and maintenance requirements. This enables proactive maintenance planning, minimizing downtime, and optimizing operational efficiency.

6. Credit Risk Assessment: Financial institutions can utilize the predictive model to assess credit risk and make informed decisions while approving loans or credit applications. This helps in reducing the risk of defaults and optimizing lending practices.

7. Customer Behavior Analysis: By analyzing past customer data and predicting future behavior, businesses can gain insights into customer preferences, purchasing patterns, and sentiment analysis. This information can be used for targeted marketing campaigns, personalized recommendations, and improving customer experiences.

These are just a few examples of the wide range of applications that can benefit from building predictive models using automated AI technologies. The versatility of such models allows organizations to make data-driven decisions, enhance operational efficiency, and mitigate risks across various industries and sectors.

9. Conclusion

In conclusion, this project successfully demonstrates the development of a predictive system using IBM AutoAI to address the issue of fraudulent transactions. By automating the process of building predictive models, the system enables efficient and accurate predictions, ultimately leading to reduced monetary loss and enhanced risk mitigation.

The web application developed as part of this project serves as a user-friendly interface for accessing and utilizing the predictive capabilities. It leverages a dataset comprising various factors such as gender, marital status, dependents, education, self-employment, applicant income, co-applicant income, loan amount, loan term, credit history, housing, and locality.

By leveraging AutoAI's capabilities, the system is able to analyze the dataset, identify relevant features, and build robust predictive models. These models can accurately predict fraudulent transactions, enabling businesses to take proactive measures to prevent financial losses and minimize risks.

Overall, this project highlights the potential of automated predictive modeling in addressing complex problems like fraud detection. It demonstrates the value of leveraging advanced tools and technologies to streamline the model-building process and improve the accuracy and efficiency of predictions. With further refinement and integration, this system can be extended to other domains and scenarios, offering significant benefits to businesses and organizations.

10. Future Scope

The above project lays a strong foundation for further exploration and enhancement. Here are some potential future scopes for the project:

1. Expansion to other domains: While the project focuses on fraudulent transactions, the developed system can be extended to other domains where predictive modeling is valuable, such as customer churn prediction, sales forecasting, or sentiment analysis. By adapting the system to different datasets and requirements, its applicability can be broadened.
2. Integration of additional data sources: To enhance the accuracy and robustness of predictions, integrating additional data sources could be beneficial. For instance, incorporating external data like social media feeds, economic indicators, or industry-specific information can provide valuable insights and improve the predictive models' performance.

3. Real-time prediction: The current system may operate on historical data, but integrating real-time data streams can enable real-time prediction and proactive decision-making. By leveraging technologies like stream processing or integrating with live data feeds, the system can provide up-to-date predictions, empowering businesses to respond swiftly to changing circumstances.
4. Model interpretability: Enhancing the interpretability of the predictive models can increase trust and understanding. Techniques such as feature importance analysis, model explanation methods (e.g., SHAP values), or rule extraction algorithms can be employed to provide insights into the factors driving the predictions, enabling stakeholders to make more informed decisions.
5. Continuous model improvement: Models can benefit from continuous improvement and adaptation to evolving patterns and trends. Implementing an automated feedback loop that incorporates new data, evaluates model performance, and triggers model retraining can help maintain model accuracy and ensure its effectiveness over time.

By pursuing these future scopes, the project can evolve into a powerful predictive system that not only addresses fraud detection but also finds applications in various domains, providing valuable insights and driving data-driven decision-making.

11. Bibliography

1. Bhattacharyya, Siddhartha, et al. "A survey of machine learning techniques for cybersecurity." ACM Computing Surveys (CSUR) 53.1 (2020): 1-45.

2. Carcillo, Fabrizio, et al. "Leveraging AI for fraud detection: A comprehensive review of the literature." IEEE Transactions on Big Data 6.4 (2020): 675-692.

3. Chen, Tianqi, and Carlos Guestrin. "XGBoost: A scalable tree boosting system." Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016.

4. IBM Watson AutoAI Documentation. Available online: https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/autoai-overview.html

5. James, Gareth, et al. An introduction to statistical learning: with applications in R. Springer, 2013.

6. OpenAI. "ChatGPT: Language Models for Task-Oriented Dialogue." OpenAI Blog, 2022. Available online: https://openai.com/blog/chatgpt

7. Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." Journal of Machine Learning Research 12.Oct (2011): 2825-2830.

8. Rashid, Nauman, et al. "Real-time fraud detection in credit card operations." Procedia Computer Science 126 (2018): 710-719.

9. Shmueli, Galit, et al. "Data mining for business analytics: concepts, techniques, and applications in R." John Wiley & Sons, 2017.

Appendix

Source code:

smartinternz02/SI-GuidedProject-524546-1688147491: Risk Prediction in Corporate Financial Management Using IBM Auto Ai Service (github.com)