
CAPSTONE PROJECT

FRAUD DETECTION MODEL

Presented By:
Dipan Basak
Institute of Engineering and Management
CSE

OUTLINE

- **Problem Statement** (Should not include solution)
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

Fraudulent transactions result in significant financial losses and increased risk for businesses. Current methods for detecting fraud are often inadequate due to the complexity and volume of transactions. There is a need for an advanced, efficient, and accurate system to predict and mitigate fraudulent activities in real-time.

PROPOSED SOLUTION

The proposed system utilises IBM's Auto AI technology within Watson Studio to automate the creation of a fraud detection model. This system will:

- Automate data pre-processing, model selection, and hyperparameter tuning
- Quickly build, evaluate, and deploy predictive models
- Provide a scalable and adaptable solution for various predictive tasks

SYSTEM APPROACH

- **IBM Watson Studio:** Collaborative environment for data analysis and model building
- **IBM Auto AI:** Automates the creation of machine learning models
- **IBM Cloud Object Storage:** Secure and efficient data storage

ALGORITHM & DEPLOYMENT

Auto AI uses multiple machine learning algorithms, including decision trees, random forests, and gradient-boosting machines.

- **Data Input:** Historical transaction data and additional relevant factors.
- **Training Process:** Auto AI trains models using historical data with techniques like cross-validation and hyperparameter tuning.
- **Prediction Process:** The trained algorithm predicts future fraudulent transactions, using real-time data inputs.
- **Deployment:** The best model is selected and deployed to generate predictions via REST API calls.

RESULT

The Auto AI-generated model demonstrated high accuracy in predicting fraudulent transactions, effectively reducing financial losses and mitigating risks. The automated process saved significant time and effort compared to manual model building, enabling faster decision-making.

SCREEN SHOTS

IBM Cloud Pak for Data

Search in your workspaces

Upgrade

4

Dipan Basak's Account

Dallas

DB

Projects / Fraud Detection / Fraud Detection

Experiment summary

Pipeline comparison

★ Rank by: Accuracy (Optimized) | Cross validation score

Relationship map ⓘ
Prediction column: Fraud_Risk

Progress map
Swap view ↺

Experiment completed ✓
16 PIPELINES GENERATED
16 pipelines generated from algorithms. See pipeline leaderboard below for more detail.
Time elapsed: 6 minutes

View log

Save code

Pipeline leaderboard ▾

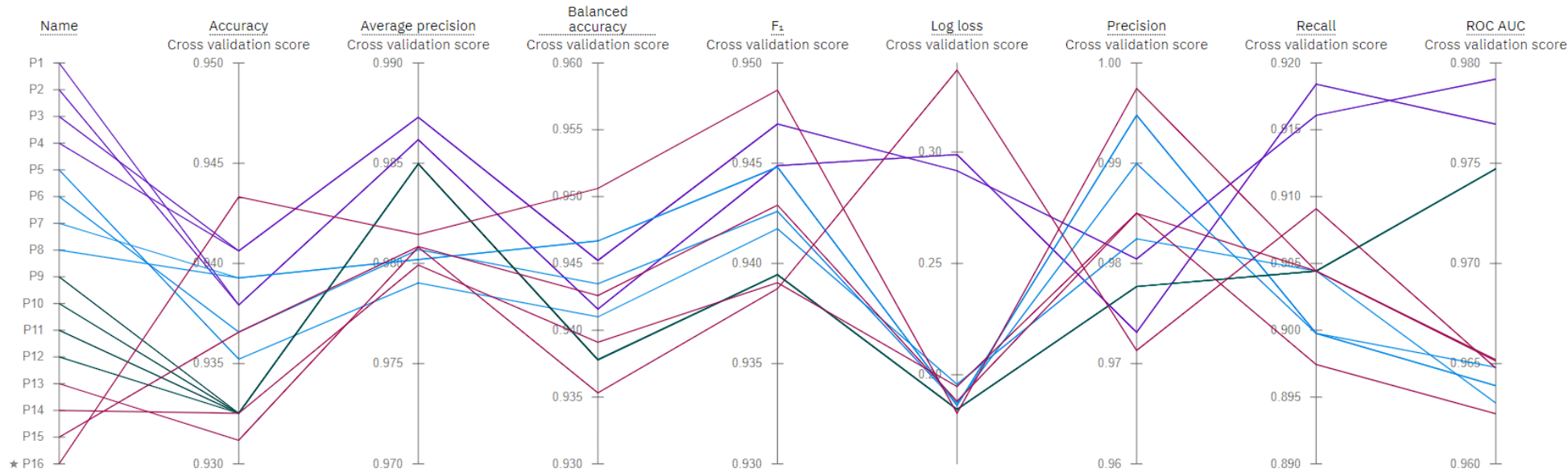
Experiment summary

Pipeline comparison

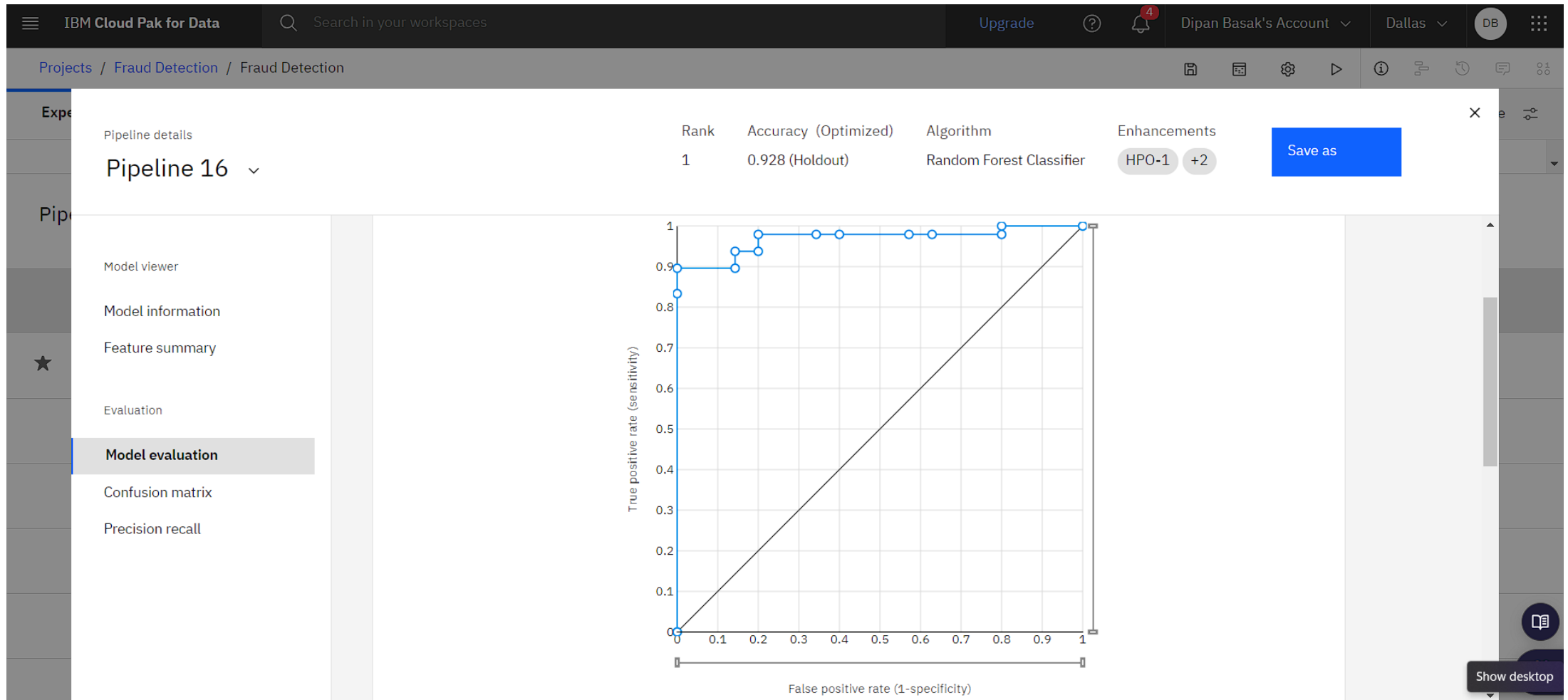
★ Rank by: Accuracy (Optimized) | Cross validation score

Metric chart

Prediction column: Fraud_Risk



Pipeline leaderboard



IBM Cloud Pak for Data

Search in your workspaces

Upgrade

4

Dipan Basak's Account

Dallas

DB

Projects / Fraud Detection / Fraud Detection

Experiments

Pipeline details

Pipeline 16

Model viewer

Model information

Feature summary

Evaluation

Model evaluation

Confusion matrix

Precision recall

Rank

1

Accuracy (Optimized)

0.928 (Holdout)

Algorithm

Random Forest Classifier

Enhancements

HPO-1

+2

Save as

Confusion matrix

Observed	Predicted		
	1	0	Percent correct
1	42	6	87.5%
0	0	35	100.0%
Percent correct	100.0%	85.4%	92.8%

Less correct

More correct

IBM Cloud Pak for Data

Search in your workspaces

Upgrade

?

5

Dipan Basak's Account

Dallas

DB

Deployments / Prediction Model / Fraud Detection /

Prediction results

Prediction type

Binary classification

Prediction percentage

1

Record

0

Display format for prediction results

☒ Table view

☐ JSON view

Show input data

	Prediction	Confidence
1	0	88%
2		
3		
4		
5		
6		
7		
8		
9		
10		

Download JSON file

CONCLUSION

- Using IBM's Auto AI within Watson Studio provides an efficient and accurate solution for fraud detection. This automation enhances productivity and ensures high-quality results, making it a valuable tool for businesses to mitigate fraud and other risks.

FUTURE SCOPE

- **Enhanced Algorithms:** Integrate advanced techniques like deep learning.
- **Real-time Processing:** Implement real-time data processing for immediate detection.
- **Broader Applications:** Adapt the system for tasks like customer behavior analysis and inventory management.
- **Scalability:** Expand to handle larger datasets and more complex scenarios.

REFERENCES

- IBM Watson Studio documentation
- IBM Auto AI overview and user guides
- Academic papers on machine learning for fraud detection
- GitHub repos for data collection

COURSE CERTIFICATE 1

In recognition of the commitment to achieve
professional excellence



Dipan Basak

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence



Issued on: 09 JUL 2024

Issued by IBM

Verify: <https://www.credly.com/go/schBYQyM>



COURSE CERTIFICATE 2

In recognition of the commitment to achieve
professional excellence



Dipan Basak

Has successfully satisfied the requirements for:

Getting Started with Data



Issued on: 09 JUL 2024

Issued by IBM

Verify: <https://www.credly.com/go/bUYCG1qF>





THANK YOU