Title: Financial Risk Prediction using IBM Auto AI

Contents

Introduction

This project discusses building a system for creating predictions that can be used in different scenarios. It focuses on predicting fraudulent transactions, which can reduce monetary loss and risk mitigation by building a web application

Using IBM AutoAI, we automate all of the tasks involved in building predictive models for different requirements. You create a model from a data set that includes the gender, married, dependents, education, self-employed, applicant income, co-applicant income, loan amount, loan term, credit history, housing, and locality.

Overview

With the increase in the digital way of life, especially when it comes to financial transactions, the risk of financial frauds cannot be ignored. A fraudulent card transaction is an unauthorised transaction that is carried out with your credit card. The thief either physically steals your credit card, or steals your card information via phishing or credit card skimming and uses it to make payments, or withdraw cash. A significant implication of a **transaction** being found to be **fraudulent** is the penalty that may be attracted in each case. Under Section 69 of the **IBC**, officers of a company which has undertaken such **fraudulent transactions** may be punishable with imprisonment for a term of up to five years and a fine extending up to rupees one crore. Fraudulent transaction is the one of the most serious threats to online security nowadays.

This project will be predicting the risk of fraud by using data which given by the user like Loan_ amount, Loan term, locality etc.

1.Purpose

The goal of the project is to develop an Fraud Risk predictor which predicts the fraud done by using certain characteristics. The model will

be trained and tested using Auto AI. The predictor is built by using Node red Application. Node red are used for creating and designing the webpage.

2.Literature Survey

Existing problem

Fraud transactions or fraudulent activities are significant issues in many industries like banking, insurance, etc. Especially for the banking industry, credit card fraud detection is a pressing issue to resolve. These industries suffer too much due to fraudulent activities towards revenue growth.

Asset-Backed Risk

Consumers borrow money for purchasing a car (auto loan), a house or running a balance on a credit card. These loans are treated as assets in the financial books of the financing entity. Bank, non-banking financial institution or a housing finance company are the financing entities.

The key value in this asset is captured by the steady stream of receivables that these assets are expected to receive. These are car loan EMIs, mortgage instalments, student loans EMI or credit card outstanding payments.

Financing institutions package these assets as asset-backed securities which are sold to institutional investors like pension funds and mutual funds.

Depending on the type of asset, profile of consumers, past record of repayment and the general nature of the economy, financial institutions offer different spreads and interest rates on these asset-backed securities. These instruments are generally treated as fixed-income securities. And hence lapped up by pension funds &

insurance companies which look for risk-averse options.

Asset-backed securities face challenges with pricing risks. This happens when investors pay more than what the bundle of assets is worth. We saw this during the 2008 financial crisis well-chronicled in Michael Lewis's book The Big Short: Inside the Doomsday Machine

Credit Risk

A credit risk happens when there is a chance that the loan borrower might skip, delay or default on his/her obligation to the bank or financial institution that has lent the money.

Let's understand this with an example.

Say you were to borrow from a bank. And they offer you the loan at a much higher interest rate of say, 13%.

This 13% interest rate consists of three parts –

1. Money that needs to be paid to the individual or institution from whom the bank borrowed the money,
2. Account for the cost of acquiring and servicing the borrower and,
3. Provision for potential delay or default in the repayment of the loan.

Management of the credit risk is the primary function of the financing institution. Underwriting the loan is a core responsibility of the lending institution. Offering a lower-than-required interest rate or offering the loan to a low-capacity-to-repay individual or firm can lead to tremendous losses and even bankruptcy.

Risk management at financial institutions

Although the loans are fixed-income in nature, financing institutions face numerous financial risks such as –

- **Interest rate risk**. An increase in interest rates will reduce the marketable value of these securities. This will tempt borrowers to refinance their loans at more favourable rates
- **Term modification risk.** Not all borrowers are able to pay per the agreed schedule. However, they are not wilful defaulters and have the intention of paying back. To ease the borrower's path to loan repayment, financiers agree to mid-term reduction in the loan's interest rate, an alternative loan type, extension in the maturity of loan or a combination of all three.
- **Pre-payment risk**. This is when borrowers pre-pay their loan. This might sound good for the financing institution but this puts them up with two problems:
    - 1. they need to find another person to loan money to and
    - 2. they have to find another investing option that earns them the same return on investment
- **Bankruptcy risk**. Borrowers may file of bankruptcy which means the loan too could be lost in its entirety to bankruptcy

Foreign Investment Risk

Foreign investment risk is the financial risk of swift and acute changes in the value of investments due to external factors like –

- Changes in accounting, reporting and auditing standards
- Nationalization. It refers to the transfer of private assets to the government often with no compensation e.g. Cuba expropriated all foreign-owned private companies after the Cuban Revolution in 1959 and Japan nationalized the Tokyo Electric Power Company

after the Fukushima Daiichi nuclear disaster

- Changes in taxation rules. The Indian government raised a tax demand of ₹11,000 crores from Vodafone. This was related to Vodafone's USD 11 billion acquisition of Hutchison Telecom in 2009. While the Supreme Court subsequently quashed the demand in 2012, the government amended its Income Tax Act retrospectively, putting the liability back on Vodafone Group. The issue is still under arbitration.
- Economic conflict. These can be trade wars like the one we see between the United States and China

Currency Risk

Currency risk or foreign exchange risk can happen when a company is having a transaction with a foreign company where one currency is stronger than the other.

The two types of currency risks are –

1. **Transaction risk**. These are losses that are likely to occur when dealing in different currency. Like the ones international food chains like Dominos and KFC face where they sell locally but report in US dollars.
2. **Economic risk**. This refers to risk associated with different political policies, varied regulations and general state of the economy in the country where business is being conducted

Interest Rate Risk

Interest rate risk is the chance that an unexpected change in interest rates will affect the value of an investment.

Here's how it works.

You purchase a bond of a housing finance company, say HDFC Limited. You bought one unit at a par-price of ₹1,000.

Because bond prices typically fall when interest rate rise, the downside interest rate risk is a drop in the value of the bond from ₹1,000 to ₹950. The reverse is true when the interest rate falls, you get a good appreciation in the value of the bond.

3.Proposed solution

Significantly reducing fraud risk will not be easy , but the available data can be used to extract the features which involves in the financial risk and try to reduce the financial fraud .Machine learning techniques can be used in the regard .The solution to the above discussed problems is to build a predictor which predicts that there is a financial fraud or risk,then we can block our credit card or block account to avoid risk and report to the bank. The following are the methodologies adopted for data analysis, prediction and display of results by integrating with the system:

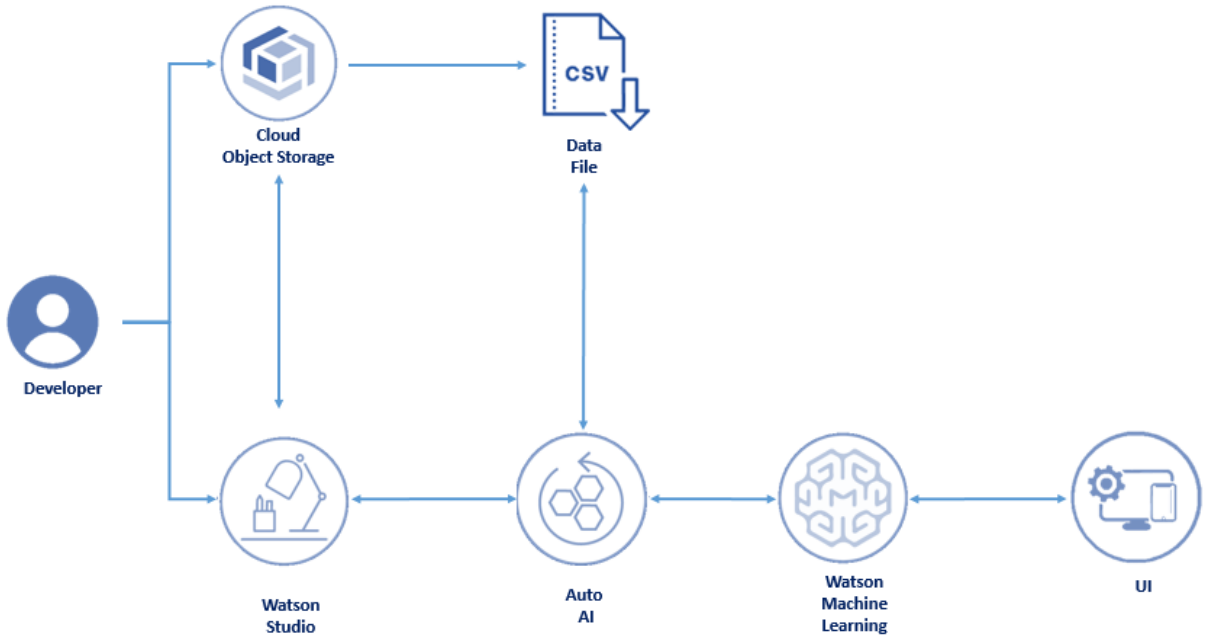
IBM Auto_AI

Node red Application

Create a model from a dataset that includes the Married, gender, Dependents, Education, Self_Employed, Applicant income, coapplicant income, loan amount,loan_term,credit housing_avaliable,Housing,Locality and fraud risk to predict the financial fraud of the individual.Using IBM AutoAI and Machine Learning algorithms, we automate all of the tasks involved in building predictive models for different requirements. Node Red application is a framework in IBM cloud that helps build web apps that could act as the interface to the user for input and output in the front and integration of values in the backend.

4.Theoretical Analysis

Block Diagram



The user creates an IBM Watson Studio Service, IBM Cloud

Object Storage Service on IBM Cloud.

● The user uploads the Financial risk data file into Watson Studio.

● The user creates an AutoAI Experiment to predict an fraud risk on Watson Studio.

● Auto AI uses Machine Learning services to create several models, and the user deploys the best performing model.

● We use the Node Red Application to connect to the deployed model and predict Risk.

5.Experimental Investigations

Various test cases were developed for each test scenario to check the correctness of the predicted value. Manual testing was also performed for each of the test case where the inputs are given and executed. The Node Red application used in the project also gives the value which is similar to the manual tested value.

Machine Learning

Machine learning can be defined as the process of teaching a computer system that allows it to make accurate predictions after the data is fed.

Classification:

Classification is a process of categorizing a given set of data into classes, It can be performed on both structured or unstructured data. The process starts with predicting the class of given data points. The classes are often referred to as target, label or categories.

The classification predictive modelling is the task of approximating the mapping function from input variables to discrete output variables. The main goal is to identify which class/category the new data will fall into.

Binary Classification:

It is a type of classification with two outcomes, for e.g. – either true or false. There are quite a few different algorithms used in binary classification in that extra tree classifier, decision tree classifier etc. We have used extra tree classifier for this model.

Extra Tree Classifier

It is a type of ensemble learning technique which aggregates the results of multiple de-correlated decision trees collected in a “forest” to output its classification result. In concept, it is very similar to a Random Forest Classifier and only differs from it in the manner of

construction of the decision trees in the forest.

Gain(S, A) = Entropy(S) - sum _{epsilon Values(A)} frac{|S_{v}|}{|S|}Entropy(S_{v})

6. Result

● Prediction using IBM Auto AI:

With Extra Tree Classifier model using gender,married, dependents,education,Self_employed,Applicantincome, coapplicant_income,variable,loanamount,loan_term,credit housing available,housing,locality and for prediction fraud risk Accuracy value in “Pipeline 7” is the top performer as the Accuracy 0.937. So the model or Pipeline 7 gives you more accuracy, which is the best model by the prediction using IBM Auto AI.

7.Advantages & Disadvantages

Advantages:

There are many advantages of predicting fraud risk. The first main advantage of the prediction is it can be used to reduce the financial fraud risk. The ability to process information almost instantly also allows the machine learning system to render an analysis in the middle of a transaction: a critical function for preventing fraud on the spot. A machine learning system, on the other hand, can handle billions of transactions and respond at lightning speed with absolute accuracy.

Disadvantages:

Having a machine learning system gets rid of the need for manual

labour. The work of 100 employees can be done in the same amount of time by a single piece of software.

It is crucial to build a proper machine learning model that doesn't malfunction. To do that, firms need trained experts who are skilled at building such systems, with an in-depth understanding of the unique field of payment fraud.

8.Applications

The prediction can be done for both classification and regression models. Being a fraud risk predictor, it can be helpful to every one. The prediction can be integrated with the company, banks, who is working to predict risk and also can be of great use to the financial fraud control purpose.

9. Conclusion

In our data analysis we have used IBM Auto AI and two classification models namely Extra Tree Classifier, Random forest Classifier for Prediction using Machine Learning Models to evaluate Financial Risk prediction data.

Prediction using IBM Auto AI:

It has been found that the Extra tree classifier model, Accuracy has obtained "Pipeline 7" as the top performer as the accuracy 0.937. Accuracy tells you how concentrated the data is around the line of best fit. So the model which gives you the more accuracy that will be taken into consideration.

10.Future Scope

The scope of the Risk predictor can be extended by adding new

features to the model and saving space and memory used to build the model. The application could also include text facility for the users to give details about the financial risk. New Machine Learning techniques can we used to predict the model easily and calculate the predict value. It may also allow software 14 applications to become accurate in predicting outcomes. It enables machines to make data-driven decisions, which is more efficient than explicitly programming to carry out certain task.

11.Bibliography

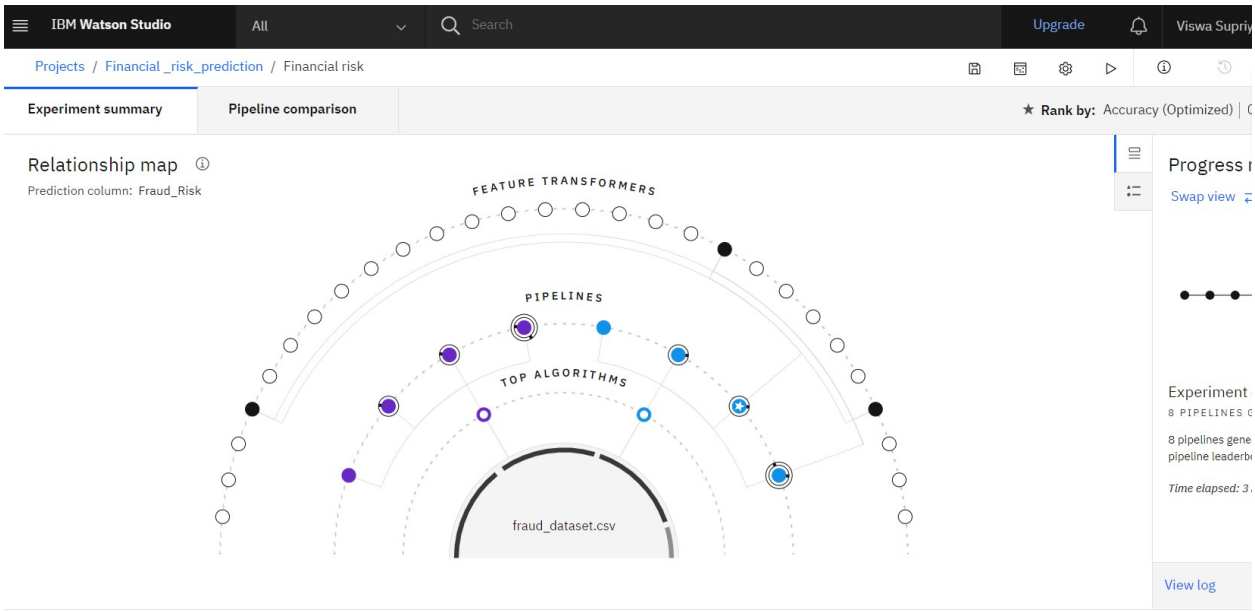
FRAUDULENT TRANSACTIONS UNDER IBC (kithandkinattorneys.in)

Fraudulent Transaction - an overview | ScienceDirect Topics

12. Appendix

Source code

- Prediction using IBM Auto AI:

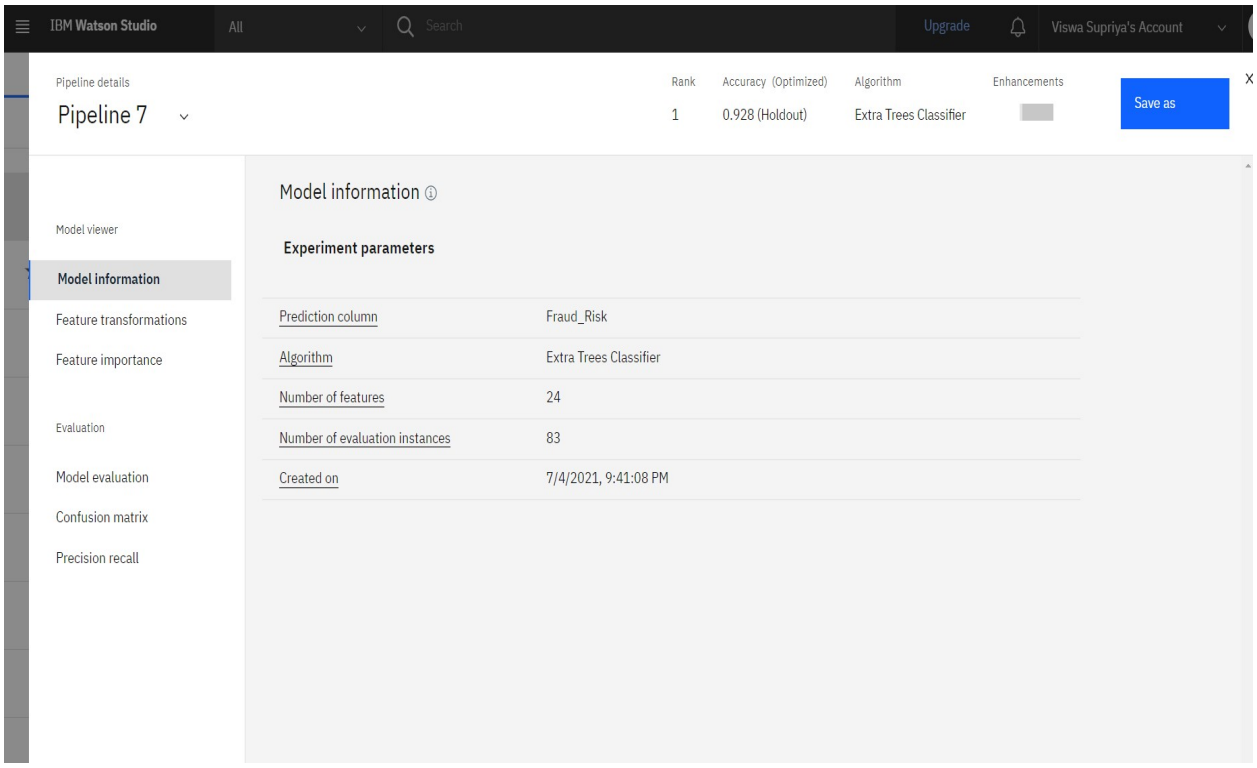


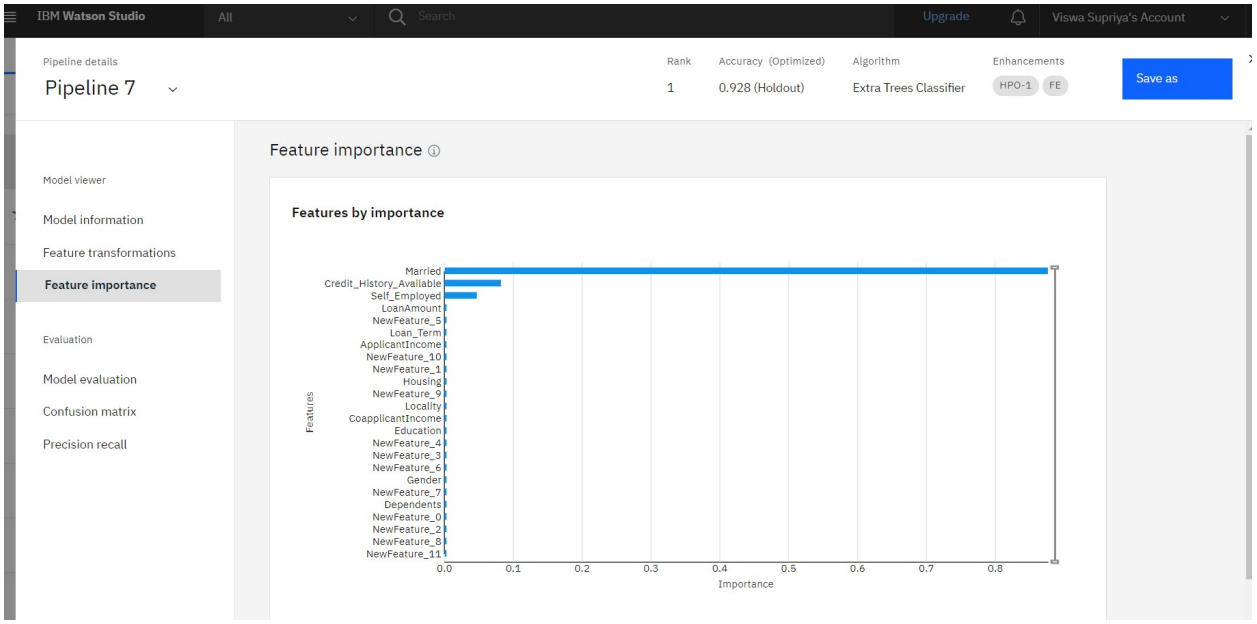
Projects / Financial _risk_prediction / Financial risk

Experiment summary | Pipeline comparison | ★ Rank by: Accuracy (Optimize

Pipeline leaderboard

	Rank ↑	Name	Algorithm	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	**Pipeline 7**	Extra Trees Classifier	**0.937**	HPO-1 FE	00:00:28
	2	**Pipeline 8**	Extra Trees Classifier	**0.937**	HPO-1 FE HPO-2	00:00:20
	3	**Pipeline 6**	Extra Trees Classifier	**0.934**	HPO-1	00:00:09
	4	**Pipeline 1**	Snap Random Forest Classifier	**0.933**	*None*	00:00:01
	5	**Pipeline 2**	Snap Random Forest Classifier	**0.933**	HPO-1	00:00:08
	6	**Pipeline 3**	Snap Random Forest Classifier	**0.933**	HPO-1 FE	00:00:25
	7	**Pipeline 4**	Snap Random Forest Classifier	**0.933**	HPO-1 FE HPO-2	00:00:14
	8	**Pipeline 5**	Extra Trees Classifier	**0.931**	*None*	00:00:02


IBM Watson Studio
All
Search
Upgrade
Viswa Supriya's Account
Pipeline details
Pipeline 7
Rank
1
Accuracy (Optimized)
0.928 (Holdout)
Algorithm
Extra Trees Classifier
Enhancements
Save as
Model viewer
Model information
Feature transformations
Feature importance
Evaluation
Model evaluation
Confusion matrix
Precision recall
Model information
Experiment parameters
Prediction column
Fraud_Risk
Algorithm
Extra Trees Classifier
Number of features
24
Number of evaluation instances
83
Created on
7/4/2021, 9:41:08 PM


IBM Watson Studio
All
Search
Upgrade
Viswa Supriya's Account
Pipeline details
Pipeline 7
Rank
1
Accuracy (Optimized)
0.928 (Holdout)
Algorithm
Extra Trees Classifier
Enhancements
HPO-1
FE
Save as
Model viewer
Model information
Feature transformations
Feature importance
Evaluation
Model evaluation
Confusion matrix
Precision recall
Feature importance
Features by importance
Married
Credit_History_Available
Self_Employed
LoanAmount
NewFeature_5
Loan_Term
ApplicantIncome
NewFeature_10
NewFeature_1
Housing
NewFeature_9
Locality
CoapplicantIncome
Education
NewFeature_4
NewFeature_3
NewFeature_6
Gender
NewFeature_7
Dependents
NewFeature_0
NewFeature_2
NewFeature_8
NewFeature_11
Features
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Importance

IBM Watson Studio All Search Upgrade

Pipeline details

Pipeline 7

Rank	Accuracy (Optimized)	Algorithm	Enhancements
1	0.928 (Holdout)	Extra Trees Classifier	HPO-1 FE

Model viewer

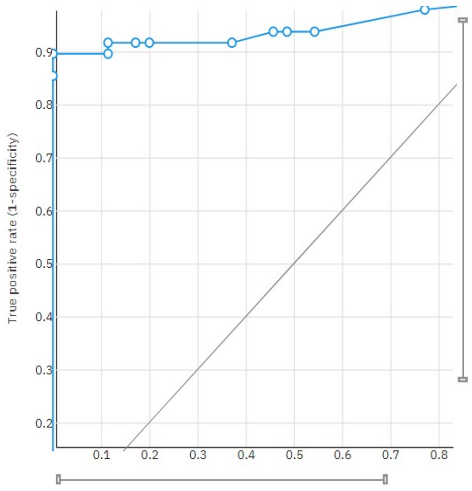
- Model information
- Feature transformations
- Feature importance

Evaluation

- Model evaluation
- Confusion matrix
- Precision recall

Model evaluation

ROC curve



IBM Watson Studio All Search Upgrade

Deployments / Financial Risk

Add to space +

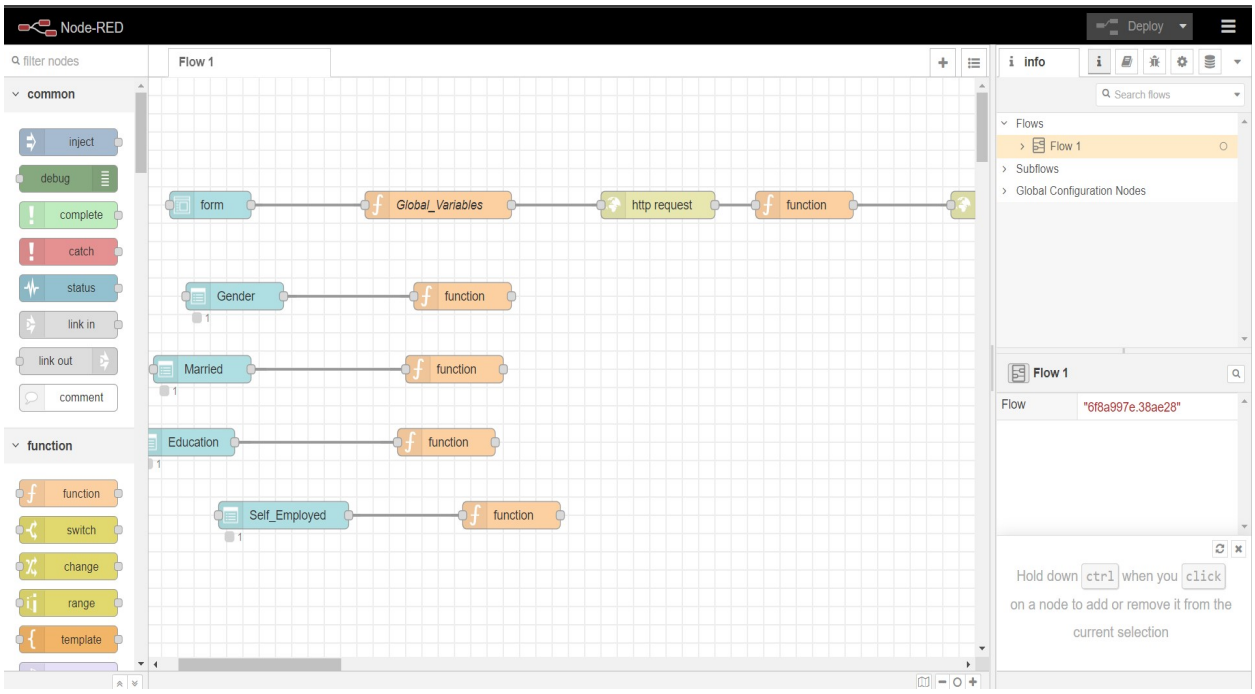
Financial Risk

Assets | **Deployments** | Jobs | Manage

What deployments are you looking for?

Deployments (1)

Name	Type	Status	Asset	Tags	Last modified
Financial Risk Prediction	Online	Deployed	Financial risk - P7 Extra Trees Classifier		Jul 4, 2021 9:45 PM



prediction

Enter the values

Gender Male

Married Married

Education Educated

Self_Employed Employed

Dependents *
1

ApplicantIncome *
4583

CoapplicantIncome *
1508

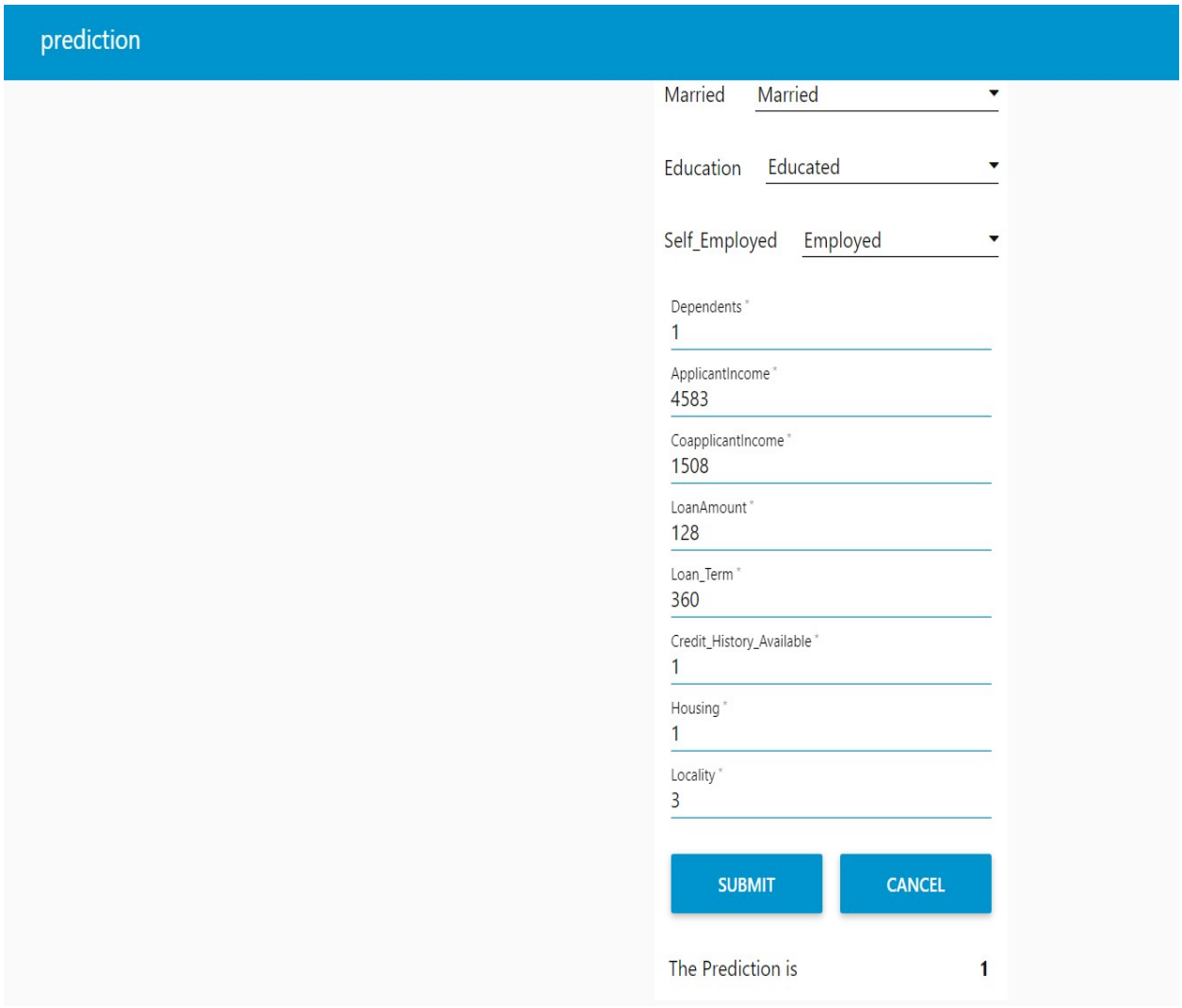
LoanAmount *
128

Loan_Term *
360

Credit_History_Available *
1

Housing *
1

Locality *
3


prediction
Married
Married
Education
Educated
Self_Employed
Employed
Dependents *
1
ApplicantIncome *
4583
CoapplicantIncome *
1508
LoanAmount *
128
Loan_Term *
360
Credit_History_Available *
1
Housing *
1
Locality *
3
SUBMIT
CANCEL
The Prediction is
1