IBM AI MODEL(MOD2)

Module 2

Prepare your machine learning project

Your AI machine learning project

Imagine that you’re an up-and-coming coder looking for your big break. Yesterday, you got a phone call from a major bank in Germany. A few hours later you found yourself flying across the ocean on a transcontinental jet! This could be your chance to make your name in AI machine learning. You leaned your seat forward and studied the materials that the bank sent you.

A terrific opportunity

The bank wants to protect itself from the risk of making loans that are never repaid.

- It’s preparing to launch a new website on which customers can submit requests for loans up to €10,000 and, if their credit is sound, get instant approval.
- The good news is that this will attract customers who want to apply for loans quickly, without a lot of fuss.

- The bad news is that it will give the bank almost no time to investigate each applicant’s credit risk, then reject those who might default on their loan, ultimately costing the bank millions of euros.

The bank wants you to lead the development of an AI machine learning system that can **rapidly predict the risk of granting a loan** for each customer who applies.

The scope of your experiment

Using simulations of IBM Watson Studio, you’ll build an AI machine learning model, train algorithms on the same data set of bank loans and defaults, then test them competitively, using additional data, to see which predicts defaults most accurately.

Your data set

- You’ll use data from a real German bank that's fictionalized for this simulation.
- The file contains records of five thousand past loans.
- Each record includes ample data such as the customer’s credit rating, bank account balances, loan purpose, and more—and whether the borrower paid or defaulted on the loan.

Your to-do list

Using IBM Watson Studio and a data set, you will simulate the following steps.

1. **Set up a new project** in IBM Watson Studio.
2. **Create a learning model** designed to predict whether an applicant is more likely to pay back or default on a loan.
3. **Train your model** on 90% of your data set, including loan outcomes.
4. **Test the model** on a test set of data (the remaining 10% of the data in the data set) to see which algorithm gives the best predictions.
5. **Save the final model** with highest confidence as a Jupyter notebook.

Important note

You do not need to download the spreadsheet of data, access IBM Cloud, or work in IBM Watson Studio to complete this course. You’ll use the following simulations to practice the steps and learn in a safe, pretend environment.

Let's get started! Today, you find yourself in the offices of the German bank, with several of their IT specialists looking over your shoulder. It’s time to show them how you work!

Simulation: Start your project and upload data

In this simulation, you’re going to use IBM Watson Studio to perform the steps to start a new project, then upload the data set of past loans and their outcomes. There are **27 steps** to complete.

> PDF is uploaded in the github file naming test case setup

Good work so far!

You’ve completed your initial setup and are ready to create the actual experiment.

- You provisioned **IBM Watson Studio** as a service.
- You created a new **project**.
- You uploaded the **data** about past loans and defaults that you’ll use for machine learning training.