

# **Configuration Manual**

MSc Research Project AI for Business - MSCAIBUS

Tural Abdullayev Student ID: X23140399

School of Computing National College of Ireland

Supervisor: Dr. Muslim Jameel Syed

#### National College of Ireland MSc Project



#### Submission SheetSchool of Computing

Student Name:	Tural Abdullayev		
Student ID:	X23140399		
Programme:	MSCAIBUS	Year: 2024	
Module:	MSc Research Project		
Lecturer:	Dr. Muslim Jameel Syed		
Date:	12/08/2024		
Project Title:	Utilizing Machine Learning to Detect Diabetes Risk		
Word Count:	6 Page Count: 907		

"I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project."

Signature: Tural Abdullayev

**Date:** 12/08/2024

#### PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiplecopies)	
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

## **Configuration Manual**

Tural Abdullayev Student ID: X23140399

### **1** Introduction

Diabetes or diabetes mellitus is a health condition resulting from an increment in blood clucose levels (IDF, 2019). In the 21<sup>st</sup> century, diabetes is considered one of the fastest-growing health concerns globally, considering that the number of people with diabetes over the past two decades has tripled. This is because of increased new cases and prevalence of diabetes. The International Diabetes Federation (IDF) indicates that the number of people infected with diabetes in 2019 was approximately 463 million. According to IDF, if the situation is not dealt with, approximately 580 million people will have diabetes by 2030, and this will increase to over 700 million by 2045 (IDF, 2019). Diabetes has various impacts on individuals' general health as it can lead to other complications such as damaging nerves, stroke, heart attack, leg amputation, and kidney failure and vision loss. It also has an impact on families and the economy since individuals with diabetes might be unable to work, failing to cater to their family's needs as well as support the economy through paying taxes.

According to IDF (2019), diabetes is one of the non-communicable diseases leading to disability and mortality. Various solutions have been implemented to help deal with the increased cases of diabetes, such as insulin, which helps manage glucose levels. Although this is the case, there has been a need to incorporate technology to help manage diabetes. According to the World Health Organization (WHO), no solution can be found to manage and monitor diabetes (WHO, 2017). Although this is the case, smart devices such as smartwatches have effectively predicted blood glucose levels, detected risk events early, and effectively adjusted insulin doses. This can also help increase the patient's quality of life. Due to the changing technology, there is room for improvement to ensure that smartwatches and other AI-generated solutions effectively manage glucose levels and diabetes.

### 2 Hardware Requirements

The solution's two components are the machine-learning module and the CASB service. Oracle cloud infrastructure is used to deploy the CASB service. On the local computer, the machine-learning model was constructed using the settings shown below:

- 64-bit operating system, x64-based processor
- Processor: 11th Gen Intel(R) Core(TM) i7-1195G7 @ 2.90GHz 2.92 GHz00GHz
- RAM: 8.00 GB (7.65 GB usable)
- Storage: 512 GB

Having at least 4GB of RAM and a dual-core Intel i5 processor or something comparable is recommended to execute the code efficiently and use CASB without experiencing CPU overhead.

### **3** Software requirements

This project utilizes Python and Java programming languages. The Apache Flink task, which uses the trained model to detect incoming events, is written in Java. Java is also utilized for data aggregation, database updates, and processing incoming events from the auth layer.

### 3.1 Development Environment

The machine learning code runs on a schedule, trains the model, and is written in Python. Jupyter Notebook is the IDE that was utilized to run the PCA model's Python code. Install Jupiter Navigator in order to construct the Python platform. Jupiter includes all Python programming languages' integrated development environments (IDEs), including Jupyter Notebook, JupyterLab, and Python 3.



#### **Figure 1: Jupiter Interface**

### 3.2 Libraries

Numerous Python libraries are utilized, including pandas, pyhocon, numpy, and scikit-learn, which are

installed and imported as indicated below.

```
# Import libraries
import numpy as np
import pandas as pd
%matplotlib inline
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')
from sklearn.utils import shuffle
from sklearn import neighbors, datasets, preprocessing
from sklearn.model selection import train test split
from sklearn.metrics import accuracy_score
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression
from sklearn.neighbors import KNeighborsClassifier
from sklearn.ensemble import RandomForestClassifier
# Read the dataset
df = pd.read_csv('dataset_work_final.csv', sep=",")
df
```

#### **Figure 2: Libraries**

Pandas - a tool for manipulating and analyzing data

Numpy - used to enable arrays

Sklearn - offers a range of statistical analysis tools and machine learning models.

Pickle - is a program that transforms Python objects into byte streams, allowing the reading and writing of data to and from databases.

Logging - makes it possible to write event statuses to log files.

### 3.3 Cloud Environment

This study employed Oracle cloud infrastructure (OCI). The OCI creates a cloud tenancy, and the CASB solution monitors all of the tenancy's resources. However, all cloud service providers may utilize this solution, and it can be combined with any current CASB system.

### References

Jupyter Packages Repository: Explore available packages on PyPI: <u>Jupyter Packages on PyPI</u>
Principal
Component
Analysis:<u>https://github.com/jakevdp/PythonDataScienceHandbook/blob/master/notebooks/05.09Principal-Component-Analysis.ipynb</u>

Installation Guide for Jupyter: Official Jupyter installation guide: https://jupyter.org/install