



## Challenge 1

### Technical Document Specification

## Table of Contents

1. Problem Overview .....	3
2. Form of Challenge .....	3
3. Essential Skills and Knowledge .....	4
4. Challenge 1 Description.....	4
5. Challenge 1 Datasets .....	4
6. Challenge 1 Marking Rubric.....	5
7. Challenge 1 Submission .....	5

# Challenge 1

## Technical Document Specification

---

### 1. Problem Overview

South Africa is a country with a high youth unemployment rate, which is currently at 46.3%. Youth unemployment and under-employment are pressing challenges in developing countries, serving as crucial indicators of economic health. South Africa is not spared from these challenges. As of the first quarter of 2023, South Africa's youth unemployment rate, focusing on jobseekers between the ages of 15 and 24, rose to 62.1%, marking the highest rate in a year. The total number of unemployed youth (ages 15-34) increased by 241,000, reaching 4.9 million. These statistics show that youth unemployment is a serious problem, as it can lead to poverty, crime, and social unrest. There are a number of factors that contribute to youth unemployment in South Africa.

One factor is the lack of skills and training among young people. Many young people do not have the skills that are required by employers, which makes it difficult for them to find jobs. Another factor is the lack of job opportunities. The South African economy is not growing fast enough to create enough jobs for all of the young people who are entering the workforce. The consequences of youth unemployment are severe.

Unemployed young people are more likely to live in poverty. They are also more likely to commit crimes, as they may feel that they have no other options. Additionally, youth unemployment can lead to social unrest, as unemployed young people may feel that they are being marginalised by society. Several things can be done to address youth unemployment in South Africa.

One important step is to improve the quality of education and training that young people receive. This will help to ensure that young people have the skills that they need to find jobs. Additionally, the government needs to create more job opportunities for young people. This can be done by investing in infrastructure and by supporting small businesses. It is also important to address the root causes of youth unemployment, such as poverty and inequality. This will require a long-term approach, but it is essential to ensure that all young people in South Africa have the opportunity to succeed.

### 2. Form of Challenge

Challenge 1 is a group project.

### 3. Essential Skills and Knowledge

The following skills and knowledge are required:

- Proficiency in Pre-Processing, involving the preparation and cleansing of data.
- Expertise in Statistical Analysis to extract meaningful insights from datasets.
- Adeptness in Data Visualization for effective communication of findings.
- Competence in Machine Learning to build robust forecasting models.

Participants have the flexibility to utilise any language, framework, development environment, library, or API in their forecasting and modelling endeavours.

### 4. Challenge 1 Description

Successfully forecasting youth unemployment rates and predicting individual employment statuses necessitates a comprehensive strategy integrating data analysis, economic comprehension, and social considerations.

Participants are tasked with precision in forecasting regional unemployment rates based on the provided dataset. Moreover, they are required to utilise historical data to predict individual employment statuses. Emphasis should be placed on the models' predictions and pivotal features affecting the predicted unemployment rate and individual employment status.

### 5. Challenge 1 Datasets

The datasets required for this challenge can be downloaded at:  
[https://drive.google.com/drive/folders/1xPzKpepPcbepxOGUHmuNt3KmT1hcJp4?usp=drive\\_link](https://drive.google.com/drive/folders/1xPzKpepPcbepxOGUHmuNt3KmT1hcJp4?usp=drive_link)

There are two types of datasets available on the google drive that need to be used for forecasting and for the prediction.

#### **Forecasting:**

Source of dataset - <https://data.worldbank.org/indicator/SL.UEM.1524.ZS>

Dataset - API\_SL.UEM.1524.ZS\_DS2\_en\_csv\_v2\_5996635

### Predict employment status:

Source of dataset - <https://www.datafirst.uct.ac.za/>

Dataset - ghs-2012-person-1.0 to ghs-2022-person-v1

## 6. Challenge 1 Marking Rubric

You will be assessed on the following elements:

- Data understanding and exploration.
- Classification model.
- Forecasting model.
- Documentation.
- Bonus points

## 7. Challenge 1 Submission

- You are required to submit your code by **23:59pm on 29 November 2023**.
- You are required to record a 15 minutes presentation of the work you have done.
- Your submission must be done on the NextCloud platform - <https://sdcdata.dirisa.ac.za/index.php/login> (More information is shared on slack and via the information pack).
- Only **one** team member needs to upload the relevant submission that comprises of code and documentation.
- The team member needs to create a shareable link.
- This link must be shared with Dr Nobubele Shozi via Slack.