Here is the background information on your task

**Welcome to Digital Intelligence!**

Imagine you had advanced business analytics to predict every interaction with customers, every moving part in your supply chain and every financial transaction. Imagine you could react to events before they happen. This is all possible with the data you already own.

PwC turns our clients into a data-driven organisations, working with them to define an AI strategy for optimised data analysis and customer interaction. As the power of data and AI is enormous, PwC acts with utmost care to use it responsibly, integrating risk mitigation while translating ethical concerns into algorithms and data sets from the start.

Initially, we would like to emphasise that our Job Simulation is based on projects one could typically experience during the first months or years with us. It is neither expected nor necessary to be able to perform such tasks without any guidance and support to start working with us. We know our tasks are diverse and you're maybe an expert only in a few of these areas, but give it a try, this experience should be a chance to broaden your horizon.

So level up with us, and use this chance to deep dive into our business and discover a career in Digital Intelligence.

**Task 2: Responsible AI**

Task Overview

**What you'll learn**

* The significance of Responsible AI in decision-making.
* How to build a Python classification model.
* Analyzing feature importance in a classification model.

**What you'll do**

* Build a Python classification model.
* Analyze which features affect the model's decisions the most, both overall and for specific cases.
* Create visual representations and provide comments on your findings.
* Here is the background information on your task
* Responsible AI is top of mind for business leaders, along with an urgent need to understand the decision process behind these algorithms. With great potential comes great risk. Not being able to understand an AI system can lead to a “black box” effect, which limits an organisation’s ability to explain and defend business-critical decisions.
* PwC’s Responsible AI approach can help. We provide services to help explain both overall decision-making and also individual choices and predictions, tailored to the perspectives of different stakeholders.
* Jakob, your team manager, is keen to meet you. Ready to hit the road? Our first stop is Geneva. Here’s an overview of the project and what I need you to do.
* A large bank has asked us to evaluate the marketing algorithms they use for retail banking. Their sophisticated phone marketing algorithm predicts whether a certain person will subscribe to a term deposit or not. Based on that assessment, the bank then optimises its phone calling strategy. With this algorithm, the bank has been successful in predicting which clients are more likely to subscribe to their term deposits.
* Management is now interested in finding out how a classification model can lead to certain decision-making processes.

Here is your task

**Ready? Download Jupyer Notebook and start creating. For your additional analysis and comments, you could use a PowerPoint slide deck and as well or do everything in the tool.**

* Create a classification model in Python. Feel free to use libraries such as SkLearn, Keras/Tensorflow or Pytorch.
* Analyse the model and provide insight into which features/variables influence the outcome of the classification the most: on a global level, and specifically for observation #4 and #20.
* Develop some form of plot or graph with brief comments on your observations.

Heads up! Jakob left you a voice message to help you with this task.

*“Hi there,*

*Thanks for helping with the task. And don’t worry too much about the performance of the model. What’s important is that you create a classification model and that you try to evaluate the relevance of potential features, both on a global and a local level.*

*Getting familiar with the Lime and SHAP libraries will help.*

*Thanks and catch you later.”*

This experience is self-paced. However, Jakob is meeting the management in 1.5 hours and needs your great model and analysis. We recommend you spend no more than that completing this task.

Research and methodology: PwC’s Global Artificial Intelligence Study

**Economic impact of AI by 2030: Net effect of AI, not growth prediction**

Our results are generated using a large scale dynamic economic model of the global economy. The model is built on the Global Trade Analysis Project (GTAP) database. GTAP provides detail on the size of different economic sectors (57 in total) and how they trade with each other through their supply chains. It gives this detail on a consistent basis for 140 different countries.

When considering the results, there are two important factors that you should take into account:

1. Our results show the economic impact of AI only – our results may not show up directly into future economic growth figures, as there will be many positive or negative forces that either amplify or cancel out the potential effects of AI (e.g. shifts in global trade policy, financial booms and busts, major commodity price changes, geopolitical shocks etc.).
2. Our economic model results are compared to a baseline of long-term steady state economic growth. The baseline is constructed from three key elements: population growth, growth in the capital stock and technological change. The assumed baseline rate of technological change is based on average historical trends. It’s very difficult to separate out how far AI will just help economies to achieve long-term average growth rates (implying the contribution from existing technologies phase out over time) or simply be additional to historical average growth rates (given that these will have factored in major technological advances of earlier periods).

These two factors mean that our results should be interpreted as the potential ‘size of the economic prize’ associated with AI, as opposed to direct estimates of future economic growth.

**AI Impact Index**

Our sector specialists worked with market participants and our partners at Fraunhofer to identify and evaluate use cases of AI across five criteria:

* Potential to enhance personalisation.
* Potential to enhance quality (utility value).
* Potential to enhance consistency.
* Potential to save time for consumers.
* Availability of data to make these gains possible.

Specific scoring parameters were derived for each criterion, and scores range from 1-5 (1 being lowest impact, 5 being highest). The parameters were weighted to arrive at a total Potential AI Consumption Impact. We also evaluated technological feasibility, and other drivers and inhibitors of consumer uptake. The results helped us to gauge time to adoption, potential barriers and how they can be overcome.

Here is your task

You have time now. It’ll take us three hours to get there.

**As part of the quantitative finance team, you and Jakob prepare a PowerPoint document containing the results of your portfolio valuation. If needed, provide your code in your slide deck to explain how you arrived at your solution.**

As you are supporting the audit colleagues with the audit for the year 2020, the value of the portfolio shall be determined as of 31 December 2020. For the valuation, you are free to use a tool of your choice (e.g. R, Python, Excel). The following steps will guide you through the valuation:

1. Inspect the historical data provided by the client. The data ranges from June 2019 until December 2020. Every month constitutes a vintage and the data includes the loan amount that was originated per vintage, as well as the repayments that have been observed up until and including December 2020 (the vintages are given as rows and the columns specify the period of the repayment).
2. Based on the provided data, compute the historical repayment percentages, i.e. every repayment’s share of the origination amount.
3. Compute the expected repayment percentages for all vintages over the lifetime of the loans. Details on how the expected repayment percentages are to be computed can be found in the attached assumptions PDF below.
4. From the expected repayment percentages, compute the forecasted cash flows using the origination amounts.
5. Using the assumed discount rate, derive the present value of the forecasted cash flows and of the portfolio. Don’t forget to convert the annual interest rate to a monthly interest rate.
6. The client’s estimate of the portfolio value was CHF 84’993’122.67. How close is this to your own estimate? Compute both the absolute and relative difference. Jakob tells you that the audit team considers any difference smaller than CHF 500’000 to be acceptable given the size of the portfolio. Based on the result of your valuation, conclude whether the difference to the client’s estimate falls below his threshold.

Jakob notes: The portfolio value should be rounded to two decimal places. The slides can be simple as we'll use them internally first. Focus on correct solutions to set up the team with the right insights.