|  |  |  |
| --- | --- | --- |
| **Document Instructions** | **Group Members**  **(Participated)** | **Group Members**  **(Did Not Participate)** |
| 1. Answer the questions in this document, and then save the document. 2. Rename the document file using the following format: GroupName\_GroupExMod5.docx 3. Email the renamed document to [submissions@esmelearning.com](mailto:submissions@esmelearning.com). 4. Return to the Module 5 Group Exercise page in the programme and confirm that you submitted your assignment. | **John Ade-Ojo**  **Mark Stroud**  **Martin Meier**  **Mayank Jamindar**  **Sue Mason** | **Replace this text with the names of your group members who were unable to participate.** |
| **Plagiarism Declaration** | | |
| By submitting this document, I confirm that:   1. Plagiarism consists of submitting another person's work or ideas as one's own. 2. Plagiarism is strictly prohibited. 3. This assignment is my own work, written in my own words. 4. I have not copied any part of this assignment from another person's work. 5. I have not allowed, and will not allow, anyone to submit any part of this work as their own. | | |

**Before you Begin**

Think back to the data valuation and project determination of the AI Solution that your group thought of from Module 4’s group exercise. In 100 words or less, describe the pros and cons of the AI Solution that you chose to implement for your use case within your organisation. *(This portion of the exercise will not be graded).*

Group Exercise 4

Guidelines

In this exercise, the group will weigh and determine how to handle issues such as customer buy-in, adoption (if applicable), and trust, especially with regard to data privacy, storage, transfer, and access for your proposed AI solution.

Together, look at the potential for discrimination as a result of human, data, or algorithmic bias in your proposed AI solution. Next, formulate a viable strategy to safeguard against possible discrimination at all stages of the process. Then move on to the following questions.

### Exercise

|  |
| --- |
| **Notes** |
| * Include citations and references in your response. * Make sure that the word count in your response is within 10% of the specified word count. * See the rubric in the module for information about scoring. |

1. In 500 or fewer words (excluding citations), determine how you will handle issues such as customer buy-in, adoption (if applicable), and trust, especially with regard to their data?

Our loan underwriting use case has a core challenge, the cold start problem. We need to fist attract a significant volume of users to the platform before we’re able to make use of data and benefit from network effects. It’s possible to look to form a strategy around solving this problem that addresses three key areas: Partnerships, financial health, and platforming.

Partnerships

Being a Fintech, we can operate in an agile and have the capacity to standup state of the art tech. However, brand recognition and data are not available to us on the outset. Overcoming this hurdle would require us to seek a partnership with either and incumbent or a big tech player, or another more established Fintech. Plum, a fintech that uses AI to help consumers accumulate savings partners with Monzo and Starling [1].

Financial Health

We need a solid financial proposition so that customers will be willing to share their financial data with us and partners see value in working with us. The Step Change Debt Charity estimates that 4.4 million people struggle to keep up with their credit commitments [2]. It’s evident that financial health and wellbeing is a massive opportunity and service that has large addressable market. Our proposition will look to focus on this, by providing a platform to manage credit card debt. Initially we would provide basic functionality including: a marketplace for retail credit, credit scoring, and AI enabled debt management advise. Note, the marketplace would offer refinancing and debt consolidation loans to help move customers away from high interest credit cards.

Platforming

We could attempt to position ourselves as the UK’s choice financial health platform. Taking this approach enables us to overcome the cold start problem by leveraging partnerships across smaller FinTechs, and eventually seeing the benefit of network effects taking over [3]. Our propositions could be broken out into microservices offered by partners initially via our platform. For example, a partnership with a Fintech like Cleo would enable us to embed AI driven debt and advise. Partnering with tech-savvy challenger banks like Shawbrook Bank would give us access to debt consolidation, and cheap personal loans for our marketplace. Credit scoring could initially be provided by the likes of Equifax for example. The premise here is that our platform gives us the ability to integrate seamlessly with existing Fintechs through APIs.

[1] Plum integrates chatbot solution with Monzo, Starling Bank through Open Banking – Global WealthTech Summit 2021. (n.d.). Global Wealthtech Summit. Retrieved 28 February 2022, from <https://fintech.global/globalwealthtechsummit/plum-integrates-chatbot-solution-with-monzo-starling-bank-through-open-banking/#:%7E:text=Personal%20finance%20chatbot%20company%20Plum,the%20use%20of%20Open%20Banking>

[2] Step Change Debt Charity. (2022, January). Falling behind to keep up: the credit safety net and problem debt. <https://www.stepchange.org/Portals/0/assets/credit-safety-nets/Falling-behind-to-keep-up-the-credit-safety-net-and-problem-debt-StepChange.pdf>

[3] Bahri, G. (2021, February 8). The Platformization of Banking. The Platformization of Banking. Retrieved 21 February 2022, from <https://www.linkedin.com/pulse/platformization-banking-germain-bahri/>

2. In 500 or fewer words (excluding citations), examine how you will successfully address the potential for human, data, or algorithmic bias (choose only one to discuss) to help ensure non-discrimination?

Andriy Burkov’s machine learning engineering book cites the types of bias that occur in machine learning [1].

|  |  |  |
| --- | --- | --- |
| Bias | Description | Mitigation |
| Selection Bias | Tendency to skew your data sources to those that are easily available. | Outline the reason we have chosen specific data sets and weigh up against alternative sources |
| Self-selection Bias | Bias can occur when you get data from sources that have volunteered it. | Unlikely this type of bias will negatively impact us. Although we are asking customers to “volunteer” financial data, it is data relevant to decisions we want to automate. |
| Omitted Variable Bias | Bias can occur when the model wasn’t trained on a feature that has good outcome signal. | Ensure we have identified good data sources, API are connected properly. Use as many features as possible with regularisation to help the model decide which features are important. |
| Sponsorship Bias | Bias comes from taking data sources produced by a sponsoring agent. | Careful review of data sources to understand who curates them and if there are any conflicts of interest before including them in any ML model. |
| Sampling Bias | Bias comes from distributions of sample data used to train ML models does not reflect the population that the model will make inferences on. | Exploratory data analysis to ensure data samples reflect inference population. Continuous model monitoring when the ML is in production to ensure training data is still relevant and to retrain the ML model if data drift is detected. |
| Stereotype Bias | Bias occurs when data used to train an ML model does not sufficiently represent a group of people. | Using under-sampling or over-sampling techniques to even out category imbalances. For credit scoring engine, ensure that financial circumstances are covered across a range of ethnic backgrounds. |
| Systematic value distortion | Bias occurs with the device making measurements or observations leading to discrepancies in training and production data for the same data features. | More often than not this type of bias occurs with unstructured data, primarily image and audio. We don’t propose using unstructured data, so this is unlikely to occur for our underwriting solution. |
| Experimenter Bias | Occurs when we interpret, search for, favor or recall data or information in a way that affirms one’s prior belief system. | Ensuring that proper data selection procedures are in place for curating ML data. Having a second pair of eyes to validate data selection. |
| Labeling Bias | Occurs when labels are assigned to unlabeled data through biased processes. | Ensuring consistent definitions for labels. For example, in loan underwriting using banking standards to label defaulted loans (IFRS 9). |

Additionally, we can avoid algorithmic bias by ensuring are machine learning models for underwriting are developed following best practices. I.e., using a training, validation and test set properly to evaluate model performance, model monitoring and re-training loops, and setting transparent model performance criteria.

[1] Burkov, A. (2020). Machine Learning Engineering. True Positive Inc.

3. In 500 words or fewer (excluding citations), analyse what concerns might be raised by your legal department regarding customer privacy, and how will you ensure compliance with current—and any anticipated future—regulations.

4. In 500 words for fewer (excluding citations), examine what ethical standards must you ensure your project meets, and what best practices will you follow, to assuage concerns and build customer trust?

**Post-Exercise Reflection**

Now that you have completed the exercise, think back to the data valuation and project determination for your AI solution that your group identified in Module 4’s group exercise. In 100-200 words, describe how the exercise you have just completed modified your idea. *(This portion of the exercise will not be graded).*