Capstone Project Proposal

Lejaah Alowayidh

Business Goals

Project Overview and Goal What is the industry problem you are trying to solve? Why use ML/AI in solving this task? Be as specific as you can when describing how ML/AI can provide value. For example, if you're labeling images, how will this help the business?	HR Process Automation Repetitive, low value tasks take time, more employees, and employees tend to make mistakes more often while working on them. That's why the product aims to automate and accelerate the pace of HR processes. Repetitive, low value tasks take time, more employees, and employees tend to make mistakes more often wile working on them ML/AI could do that when included in: - recruiting (scan CV''s). - chatbot (for both employees and HR employees). - analyze and report HR metrics (employee's performance, employee's satisfaction, turnovers, budget, training…etc.
Business Case Why is this an important problem to solve? Make a case for building this product in terms of its impact on recurring revenue, market share, customer happiness and/or other drivers of business success.	HR metrics are vital and have a strong impact on business success. Especially the strategic ones. Automating repetitive, low value tasks allows HR employees to focus on more strategic, creative work. Which eventually lead to minimizing turnover, recruiting and onboarding coasts, also increasing employee satisfaction which increases their productivity in return…etc.
Application of ML/AI What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?	ML/AI will be used to accomplish many tasks at HR department. It'll be used in recruiting to scan CV's and recommend the best matches based on keywords. in a chatbot to provide instant response. in analyzing HR metrics and providing suggestions to improve them.

Success Metrics

Success Metrics What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you will establish a baseline value to provide a point of comparison.	- Submit to interview ratio, Retention rate, Employee happiness, Training efficiency, Chatbot goal completion rate. - And Talent turnover rate, Time to hire, Chatbot fallback rate. What we want is to increase the first group and decreasing the second, that's how we know how success the product is.

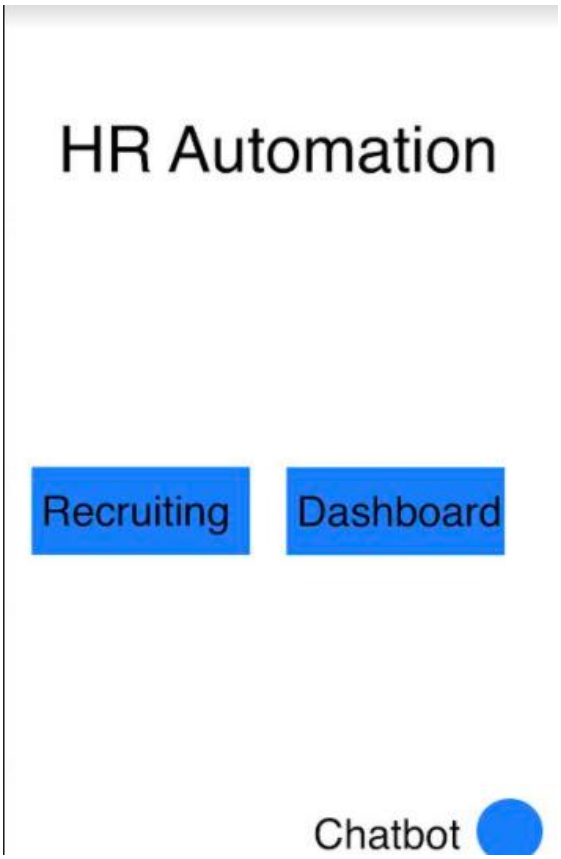
Data

Data Acquisition Where will you source your data from? What is the cost to acquire these data? Are there any personally identifying information (PII) or data sensitivity issues you will need to overcome? Will data become available on an ongoing basis, or will you acquire a large batch of data that will need to be refreshed?	The product nature requires collecting data from the organization itself, business record, experts' professional opinion. Therefore, this process would be simple, affordable, no need for PII, and most of it requires a large batch of data that will need to be refreshed, while analyzing HR metrics requires an ongoing data.
Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	Data will be collected from the organization itself, business record, experts' professional opinion. That's where bias could appear, data accuracy depends on the size of the company; the bigger, the accurate. However, small companies can still benefit if we use component's records.
Choice of Data Labels What labels did you decide to add to your data? And why did you decide on these labels versus any other option?	Recruiting: simple binary classification system, "Fit" and "Reject" and the reason is we need this process to be simple as possible. However, we could optimize and add more features like ranking the candidates. Other data can't be labeled, it can be either statics or natural language processing.

Model

Model Building How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team, and why?	Using an external platform to build my model would be much easier. Although, it can be compromising due to their terms of service and giving them access to my data. However, the model is simple and contain sensitive information. Therefore, choosing in-house data science team is the appropriate decision here.
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	Accuracy, Precision, Recall, F1 Score can be used to measure the model performance. However, it would be best if we can get results closer to 100% as possible.

Minimum Viable Product (MVP)

Design What does your minimum viable product look like? Include sketches of your product.	
Use Cases What persona are you designing for? Can you describe the major epic-level use cases your product addresses? How will users access this product?	The persona I'm designing for is HR employees in all different industries. The model provides an HR Dashboard, CV scanner, and chatbot, it should be available to all HR employees in the organization.
Roll-out How will this be adopted? What does the go-to-market plan look like?	The pre-lunch plan includes comparing with component's product, and testing. While post-lunch plan includes constant updates, and taking suggestion from HR employees.

Post-MVP-Deployment

Designing for Longevity How might you improve your product in the long-term? How might real-world data be different from the training data? How will your product learn from new data? How might you employ A/B testing to improve your product?	Since we're using from the organization itself, real world data won't be a big issue. However, to improve the product we must provide new data continuously, this could be every quarter or every year. We could also employ A/B testing by deploying new updates to few branches, and modify it based on the feedback or publish it, if it did work perfectly.
Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	The product should be monitored on a frequent basis by our team, and keep feeding the system more diverse data, and testing.