Capstone Project Proposal

<*Tamer Abdelaty*>

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?	The problem that I will try to solve is in Healthcare industry by detecting what drugs that are written by a physician to be dispensed by a pharmacist without involving in reading the prescription. So the pharmacist will operate AI system which will process the medications inside the prescription without the interference of human which will decease human mistakes. Machine Learning can help us in this project to classify and detect the medications to be dispensed inside prescriptions taking the drug index for the specific country and the specialty for each prescription for example (Pediatric , Gynaecology, GIT, etc…), we will probably uses a supervised learning algorithms for classification such as (Random Forest).
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.	The problem that I have mentioned earlier is highly important, because reading prescriptions by pharmacists is time wasting and involve in human errors which are not acceptable in healthcare industry. So it will save time of the pharmacist and the patient and ensure of perfect dispensing of the medications. Also it will save money by saving time and canceling human errors.

Application of ML/AI What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?	We will probably uses a supervised learning algorithms for classification such as (Random Forest), to classify each Medication and speciality to decide whether specialty match with medication and ensure of the correct detection of the medication inside each prescription to be dispensed. We could use and implant also text recognition technology in our model.

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.	Multiple metrics can be used. • The correctness of detection of the medication inside every prescription. • The correctness of detection of the speciality inside every prescription. • The perfect matching of the detected medications to the detected speciality

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 Data Acquisition can be done from: -The drug index for the selected & specified country. -The huge number of prescriptions to be collected and processed using text recognition technology which contain the speciality and medications for specific patient from the selected & specified country. -Prescriptions from hospitals medical records in the the selected & specified country.
Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	Algorithm Bias: The biases that can happens in algorithm, Medications have canceled by the health authority of the selected & specified country that will causes some biases Measurement Bias: Over-fitting of the model, so we think it perform well but it actually perform well in the training set only. Exclusion Bias: We may faces outliers in the data-set, which is considered as bias, we should try to handle it. Human Bias:Errors in the prescriptions which possible to be in speciality or medications.
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?	I have decide to use these labels -Specialty -Medications Because these in my opinion is the most important labels to detect the important data in prescription in the selected country for our model.

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?	I will use an in-house team, because the sensitivity of the Medical information that we collect which will force us to make it in-house.
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	There are multiple evaluating criteria can be taken • Accuracy • There are no Over-fitting It should perform very accurately to make it feasible to be applied, the accuracy should be (99%) at least, also to spot the over-fitting we should look for the accuracy of the test set

Minimum Viable Product (MVP)

Design What does your minimum viable product look like? Include sketches of your product.	Portal that contains the all the speciality medications present in the selected country with text recognition system that uses our ML model efficiently.

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?	Pharmacists will use this system on web page to dispense patients medications without interfering.
Roll-out How will this be adopted? What does the go-to-market plan look like?	1. Who: The targeted people is the Pharmacists. 2. .What: Machine Learning model to detect the correct medicine and correct specialty of the patients prescriptions. 3. Why: to save the time and ensure of the accuracy of the medications dispensing process. 4. Where: in Pharmacies and dispensing facilities 5. How: Through a web-page. 6. When: After testing and deploying the model, and make sure it works correctly.

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?	We will use large amount of data, from multiple sources and places will enhances the model performance, there are a margin of error for example (medications in a prescription may contain canceled drug by health authority or wrong specialty. So that our model it will learn from applying large amount of data on it using real-world data, I can use A/B testing to decide if the canceled drug exist in the update drug index database so it will-not affect the model accuracy. So the null hypothesis will be using medications in the model, and the alternative hypothesis will be not using the medications in the ML model.

Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	I can monitor the canceled drugs using updated drug index database and avoid the wrong specialty using medications to guess the correct specialty which will make our model work without bias or mitigate unwanted biases from the model in my opinion.