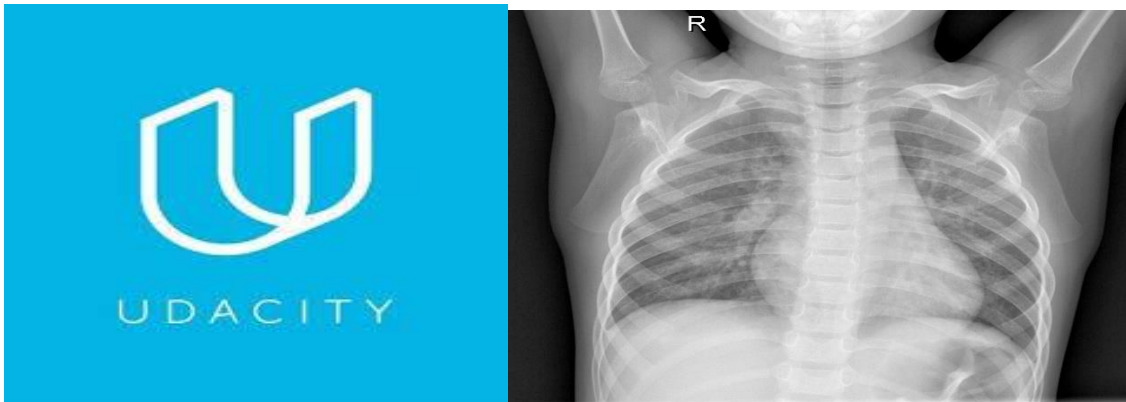


Doctor Drones

Healthcare Tracking & Monitoring ▪ From Up Above . . . Till Ground Zero
Analytical Medical Tracking From Drones to Apps



Capstone Project by: Frederick Zoreta

General Overview of Doctor Drones



Mobile & handheld app

A software that could be installed on both iPhone and Android has the ability to analyze and predict a person's well being by analyzing certain patterns. This app also has computer vision and image processing capabilities that easily identifies & classifies various chest xrays.

The product is a hybrid of both an app that functions in measuring specific characteristics of a person. Such characteristics involves:

1. An individual's temperature
2. The human kinetics aspect: changes in movements, speech patterns, and over-all physical dynamics
3. Analysis of chest xrays. Using computer vision and image classification, chest xrays would be analyzed and compared to the person's characteristics as mentioned above. For example, if a person has pneumonia, comparing these findings with the temperature and other patterns, would individual A's condition eventually lead to COVID-19? Is it SARS? Would he/she be more prone to having Ebola?
4. Utilizing ML algorithms, the app also has the ability to predict the probability of any possible reinfections; should the individual be a case.



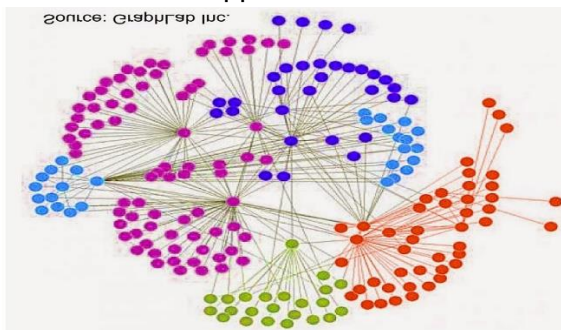
Drones: Flying Medical Assistants & MDs

The data being processed within the apps are relayed automatically in real time to the drones. Drones would then apply graph algorithms that would 'contact trace' the individual's possible contacts within a certain timeframe. It would also calculate the specific areas that the 'target individual' may have visited or inhabited.

**Me and my entire team is fully aware that such a practice may pose certain privacy and security issues in certain countries and specific jurisdictions. There would be specific configuration changes and adjustments depending on where our stakeholders & clientele would be located.

These drones would be utilizing graph algorithms and graph technology. The proposed algorithm and software would initially be compatible with 2 graph databases: Neo4j & DataStax Cassandra. Eventually it may also expand to ArangoDB and Tigergaph.

Conclusion: The specific business case for Doctor Drones is to eventually venture into other possible healthcare applications.



Capstone Project Proposal



<Name> *Frederick Zoreta* </Name>

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?

I am highly focused on the healthcare industry. This is not just to build-up on the first 2 projects, but also extremely timely regarding today's global pandemic situation.

My business (also a medical and philanthropic) goal is to create an app if an individual has COVID-19. This could be done by various ways:

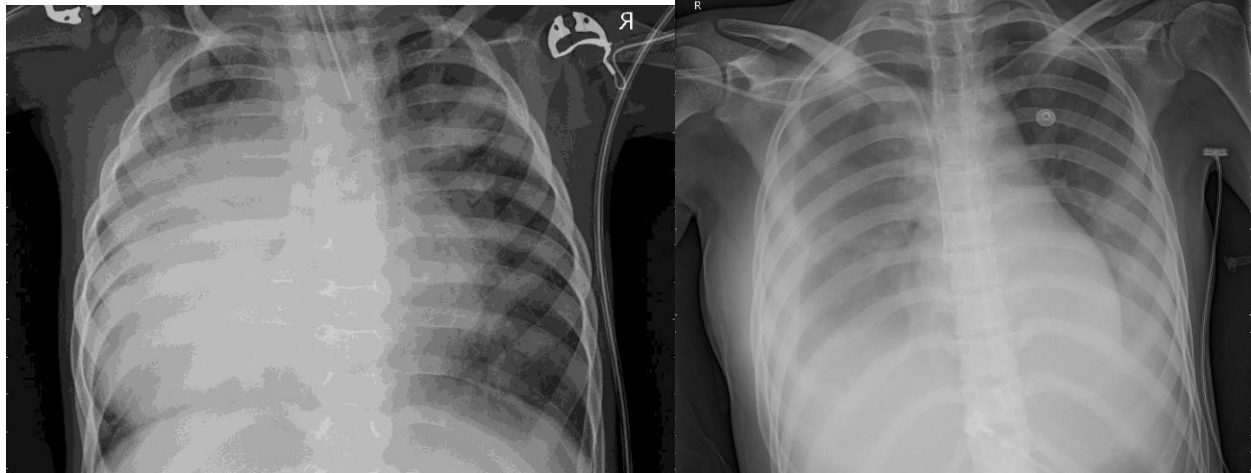
1. Measuring & analyzing the individual's temperature
2. Applying ML algorithms in determining the paths and places where the person travelled and also the people that he/she may have interacted with
Graph algorithms/ graph databases would also be utilized in this aspect (Neo4j, Cassandra, ArangoDB, TigerGraph)
3. Analyzing the movements and characteristics of the person.

This would involve analyzing the past photos (and videos if possible) of the person. This would be a good comparison if the person has sudden changes in movements, mannerisms, energy level, etc.

Image and video classification would definitely be utilized in analyzing chest Xrays and the over-all physiology.

A real life example: If a person is extremely strong and fit, with no symptoms, could he/she be just simply asymptomatic?

| | |
|---|--|
| <p>Business Case</p> <p>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.</p> | <p>My upcoming product would be of tremendous help not just in the fight against COVID-19, but it could also be utilized in determining any potential viral/bacterial threat.</p> <p>COVID-19 and pneumonia detection would simply be just the initial phase of this app/software.</p> <p>Drones would also be utilized in order to facilitate advanced contact tracing and also tracking areas of higher possible infections.</p> |
| <p>Application of ML/AI</p> <p>What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?</p> | <p>My business case would be utilizing both 'image classification', 'video analysis/ video classifications' and also 'data annotations for both images and videos'.</p> <p>On top of this, the data being gathered would be interacting with graph algorithms that are embedded within the drones.</p> <p>The main business outcome expected is mainly divided into 3:</p> <ol style="list-style-type: none"> 1. A very efficient and accurate way of detecting various respiratory diseases, with a strong focus on COVID-19. 2. Facilitate contract-tracing by analyzing the data form both xray images and videos (if applicable). This would interact with the drone's graph algorithms. The drone's would analyze the possible areas and perimeter of where and when the 'patient' have been within a certain geographical region. Drones would also act as a way of 'aerial patient watch/security guard' for certain number of positive patients. 3. Predictive analytics |



Above is a classic example of 2 pneumonia xrays. The left side is bacterial and the right side is viral. Both are non-normal, but even trained professionals within the medical imaging profession are prone to errors. If other forms of medical images are involved such a COVID-19, then the need for more AI/ML-based analysis would be highly beneficial.

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.

This type of product is extremely new to the global market. It is also unique and somewhat experimental.

Business metrics to be measured are:

1. Customer Acquisition Cost (CAC)
2. Retention Rate / Churn Rate
3. Return on Advertising Spend (ROAS)
4. Margin
5. Customer Lifetime Revenue (CLR)

Additional note:

The above success metrics were chosen for the following reasons:

- a. Our business is still a small business/ start up on its very infancy stage. It is very vital to know the upcoming financial results.
- b. As a startup, my team would probably be bootstrapping, our main focus during the early stages may be to simply focus on 'REVENUE' and 'PROFITS'.
- c. Utilizing such financial metrics may easily aid in verifying and spotting the possible financial/economical pitfalls.

1. Customer Acquisition Cost (CAC)

**** This measures acquiring a new customer. Mathematically, the fastest way to calculate this metric is to pick a 'specific time period' (time line or range) then divide the cost of marketing and sales by the number of customers you gained.**

For example (Christmas Season 2023)

Marketing Costs in US\$ / # of customers being gained.

2. Retention Rate or Churn Rate

**** This is simply the % of customers that stays with our business.**

Simply put the main goal is to keep the retention rate as high as possible. This translates as 'happy and satisfied customers' = 'Higher and constant \$\$\$\$\$\$'.

3. Return On Advertising Spend (ROAS)

**** Derived by: Sales Generated / Advertising Spending**

An example would be : if my team spends \$20,000 on advertising efforts & materials and it resulted in \$40,000. This means that the ROAS is 2. Our company generated \$2 for every \$1 that we have spent.

4. Customer Lifetime Revenue (CLR)

**** Sometimes known as Customer Lifetime Value (CLV).**

This metric is truly helpful in evaluating the quality of my team's customer service and 'public relations'. I personally believe that Customer Experience (CX) and User Experience (UX) should definitely one of the highest priorities of any business.

If a customer feels that my customer service is inadequate, then there is a high possibility that the individual may leave.

5. Margin

****Some business schools label this as 'the bottom line of everything'. It is totally vital for every business to acknowledge that we should be arriving at 'the finish line' or a 'specified target destination'.**

And this destination and finish line should be the EXACT DESTINATION that me and my team really wants to be in.

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? | <p>Data would be derived from reliable and reputable medical sources. Majority of the data would be medical images such as chest xrays. Other sources of data would be videos of an individual's movements.</p> <p>Data is highly voluntary and given with full consent from patients and stakeholders. The gathered data would then eventually kept for a certain period of time in order to be used for further analysis and to 'machine-learn' certain scenarios. This in turn would be used for predictive analytic purpose.</p> <p>The other source of data would be the 'video capture' and the specific numerical values gathered from drones. This is a sensitive matter that may impact certain privacy and security issues, depending on the jurisdiction.</p> |
| Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved? | <p>Medical images and captured videos are large data sets.</p> <p>The main bias is regarding the analysis of chest Xrays. Various Xrays being analyzed may influence the results. An example would be someone with pneumonia. Would this situation lead to Ebola? COVID? Does the ML prediction state a potential 'disease' that may not have been discovered yet?</p> <p>Will the model base the results on 'historical data'.?</p> |
| 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? | <p>Data labels would be initially limited to the following:</p> <ol style="list-style-type: none">1. Normal2. Bacterial pneumonia3. Viral pneumonia4. Ebola5. SARS6. Swine Flu7. COVID-198. High Fever <p>**Initial analysis is mostly expected to be derived from chest Xrays</p> |

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? | <p>This would normally be a case by case basis. But, due to the ultra-sensitive nature of the data being analyzed, plus the fact that the technology to be used is highly proprietary, I would prefer to hire a small team.</p> <p>My in-house team would be composed of specialist from various fields such as: ML/AI experts (4 to 6), cloud operations (either Azure or GCP), 2 mechanical or electrical engineers that specializes in drones/ aerial machines, and between 2 to 4 UX researchers /designers. These UX'ers can also act as 'product management' assistants who would work directly with me. I would also make sure that there are at least 2 of my data scientists/analysts who are highly skilled in graph algorithms and graph databases.</p> <p>I may hire a contractual genomic scientist since having a full-time genomic specialist would cost a lot.</p> <p>I would rather spend money on an inhouse team that I know I can fully trust. TRUST not just with very confidential data but also with proprietary and new technology.</p> |
| Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required? | <p>The evaluation would be based on performance metrics would be:</p> <ol style="list-style-type: none">1. Confusion Matrix2. F1 Score3. Cross Validation <p>**The first 2 have been utilized in past projects and has been proven to measure the data well.</p> <p>Number 3, cross validation is very vital in data modelling. It states that there should be a sample that should not be trained, and also test a specific sample before finalizing the model.</p> |



3 Basic User Interfaces or 'Design Mock Ups'

type something



Doctor_Drones

Healthcare monitoring from up above
and within ground zero



Username:

Password:

type somet



Welcome to Your Doctor Drones Health Tracker

Data Input

Image Classification

Video Annotation

Select Algorithms

Input Data
Lorem Ipsum
Dolor Ipsum

etc

etc





Image Analysis of Xrays




Minimum Viable Product (MVP)

| | |
|---|---|
| Design What does your minimum viable product look like? Include sketches of your product. | <p>The MVP is a simple app that functions in both iPhones and Androids. As mentioned from the summary, the app would mainly focus (initially) on using image classification among chest X-rays. Although the main focus would be COVID-19, it would also detect other respiratory illnesses.</p> <p>Drones, which is the MAIN FOCUS of 'Dr. Drones' would be deployed on a case-by-case basis. The main focus of drones would be an advanced version of graph algorithms. This is to facilitate more accurate contact tracing.</p> |
| 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? | <p>The 3 main/basic personas are attached below and also within the document. These target audience is highly focused on medical research, medical practice and analytics.</p> <p>The epic-level use cases are: extremely accurate results of findings by simply utilizing non-invasive procedures.</p> <p>The use of advanced ML/AI methods by means of drones means also faster and more accurate CONTACT-TRACING. **The drone usage is subject to each nation's & jurisdictions laws.</p> <p>Users(persona) would simply access this via the 'Dr. Drones' app that could be used either iPhone or Android.</p> |
| Roll-out How will this be adopted? What does the go-to-market plan look like? | <p>We plan to use a lot of social media marketing during the initial phase. Use of Paid Advertising (PPC) may eventually be used, but main focus is on social media marketing and utilization of professional and educational networks.</p> <p>Our go-to-market strategy is to HIGHLY FOCUS on the 3 main target audience personas. We would be utilizing YouTube as one of our advertising platforms.</p> <p>A strong focus would be placed on LinkedIn since this is where majority of professional and education-based communication happens.</p> |

3 Main Personas within the target market/audience:

Melissa : A Doctor



Age
35 to 44 years

Highest Level of Education
Professional degree (e.g. MD,

Social Networks

Industry
Health Care

Organization Size
201-500 employees

Doctor Melissa Jean

Preferred Method of Communication
Cell phone , email, professional social media account

Tools They Need to Do Their Job
Enter text here

Job Responsibilities
Initial patient care, emergency medical procedures

Their Job Is Measured By
Success is measured via patients improving over-all condition. No occurrences of being re-admitted to hospital.


Reports to
Head Doctors, MD Leads, Chief Medical Officers

Goals or Objectives
Over-All Patient Safety, (for those infected) smooth & rapid recovery from respiratory illnesses

They Gain Information By
Patient data being generated by hospital records, data form RNs (nurses) and fellow doctors.

Biggest Challenges
Over-all communication and sharing of patient data

Eric: IT Analyst/ BI Specialist/ Data Scientist



Age
35 to 44 years

Highest Level of Education
Master's degree (e.g. MA, MS,

Social Networks

Industry
Technology

Organization Size
201-500 employees

Eric aka "Mr. Analytics Pro"

Preferred Method of Communication

- Email

Tools They Need to Do Their Job

- CRM Software
- Invoicing Software
- Reporting Software
- Business Intelligence Dashboards
- Employee Scheduling Software
- Cloud-Based Storage & File Sharing Applications
- Content Management Systems

Job Responsibilities
Enter text here

Their Job Is Measured By
Accurate Analysis and Visualizations of Medical and Genomic Data Sets

Reports to
CTO, Hospital Managers, Data Science Director, Business Intelligence Director

Marcus: Bio-informatics specialist/ Genomic Researcher

Marcus "Genomics Specialist" / Medical Researcher



Age
45 to 54 years

Highest Level of Education
Doctorate (e.g. PhD, EdD)

Social Networks



Industry
Health Care

Organization Size
11-50 employees

Preferred Method of Communication

- Email

Tools They Need to Do Their Job

- Business Intelligence Dashboards
- Cloud-Based Storage & File Sharing Applications
- Content Management Systems
- Reporting Software

Job Responsibilities

Accurate findings of various activities and anomalies related to respiratory illnesses

Their Job Is Measured By

Over-all analysis of large volumes of data sets, publications of high quality journals, recognized research being presented

Reports to

CEO, CTO, Founders, Chief Scientists, Chief Medical Officers

Goals or Objectives

Highly Research Focused

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?

As time passes, we would be using the data in order to go to a certain audience / avenue where Dr. Drones would gain more attention, and hence, revenue.

Real world data may initially be very 'dirty' since there would be a variety of Xray images to be analyzed and classified. On top of this, the applicable videos of each person would also be taken into consideration, which could eventually affect the classifications models.

For the initial phase, we could use random sampling. We simply get a sample in each specific population within the target personas.

My team could also use 'experimental design'. We would be using simple and straightforward test. I would then rollout the following:

1. Our team's hypothesis
2. Randomization strategy
3. Sample size
4. Method of measurement

Monitor Bias

How do you plan to monitor or mitigate unwanted bias in your model?

Firstly, there are different types of bias in ML/AI.

These are :

1. Sampling bias
2. Measurement bias
3. Exclusion bias
4. Experimenter or observer bias
5. Prejudicial bias
6. Bandwagon effects

Our team would be following these 6 specific types of bias mitigations:

1. Identification of unwanted sources of bias
2. Identifying accurate representation of data.
3. Setting up proper rules and guidelines for eliminating bias & procedures
4. Documentation of how data is collected/analyzed and shared
5. Model evaluation for performance
6. A thorough and proper review of models that are being utilized in this project phase.

Appendix:

1. 6 Ways to reduce bias in machine learning and data analysis

<https://searchenterpriseai.techtarget.com/feature/6-ways-to-reduce-different-types-of-bias-in-machine-learning#:~:text=Six%20ways%20to%20reduce%20bias,train%20the%20machine%20learning%20model.>

2. How to detect bias in AI

<https://towardsdatascience.com/how-to-detect-bias-in-ai-872d04ce4efd>

3. 5 performance metrics every business should know

<https://www.inc.com/craig-bloem/5-key-metrics-every-early-stage-business-must-track.html>

4. A/B Testing : What role does it play in the era of Machine Learning and A.I.

<https://medium.com/capital-one-tech/the-role-of-a-b-testing-in-the-machine-learning-future-3d2ba035daeb>

5. Proven Process for developing a 'market-go-to-strategy'

<https://blog.hubspot.com/sales/gtm-strategy>