1. Idea and Proposal

Pneumonia Detection

What is the business problem?

Detecting Pneumonia fats with AI

Who are the intended stakeholders, and why is this problem relevant to them?

Doctors - it will help the doctor detect Pneumonia

Patients – Detecting Pneumonia before seeing a Specialist.

Where are the datasets available from? There are available at https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia

Which one do you like the most?

Pneumonia Detection

What type of data science approach would you use?

Binary image classifier using convolutional neural network layers.

How many rows and how many columns does the dataset have?

The dataset is organized into **3 folders (train, test, val)** and contains subfolders for each image category **(Pneumonia/Normal).** There are 5,863 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal).

2. Idea and Proposal

Natural Language Processing: Classify Amazon reviews based on the customer's ratings.

What is the business problem?

Analyzing the reviews, we will help users decide which ones to trust as well as report and adjust the fake ratings.

Who are the intended stakeholders, and why is this problem relevant to the

Marketing and customers

Where are the datasets available from? There are available at

142.8 million reviews (http://jmcauley.ucsd.edu/data/amazon) and 1.4 million answered Q&A (http://jmcauley.ucsd.edu/data/amazon/qa).

Which one do you like the most?

Pneumonia Detection

What type of data science approach would you use?

Natural Language Processing and Feature Engineering

How many rows and how many columns does the dataset have?

In this project, we use 5-core dataset of Clothing and Shoes, which is subset of the data in which all users and items have at least 5 reviews.

---

3. Idea and Proposal

Severity Detection of Knee Osteoarthritis

What is the business problem?

Detecting how severe the Knee Osteoarthritis thru X-Ray Images

Who are the intended stakeholders, and why is this problem relevant to the

Doctors - it will help the doctor read and detect Osteoarthritis

Patients – Detecting Osteoarthritis before seeing a Specialist.

Where are the datasets available from? There are available at https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity

Which one do you like the most?

Pneumonia Detection

What type of data science approach would you use?

Multiclass Classification with CNN/Keras

How many rows and how many columns does the dataset have?

The dataset contains knee X-ray data for both knee joint detection and knee KL grading. The Grade descriptions are as follows:

- Grade 0: Healthy knee image.
- Grade 1 (Doubtful): Doubtful joint narrowing with possible osteophytic lipping
- Grade 2 (Minimal): Definite presence of osteophytes and possible joint space narrowing
- Grade 3 (Moderate): Multiple osteophytes, definite joint space narrowing, with mild sclerosis.
- Grade 4 (Severe): Large osteophytes, significant joint narrowing, and severe sclerosis.