PROJECT 1: “Pain level score” based on patient notes and treatment details

1. Business scenario

A summarized patient pain level score, lets say post surgery, can help the doctor and nurse track the progress healing/any deterioration via a summary metric. Currently doctors and nurses speak to patient, take notes, analyse them in conjunction with other treatment details and hospitalization details to understand how patient is recovering. A ML model will free up a lot of time of doctor and nurses and be more accurate in guiding the doctor and hospital staff on the recovery.

2. Data options

Generate a synthetic dataset as historic data, labelled – example below:

Patient Notes, Days Admitted, Treatment/Surgery Type, Surgery Done, Pain Level

I feel a little better today,7,Liver transplant,Yes,8

The pain is unbearable,4,Physiotherapy,No,7

I can walk but it hurts,13,Medication,No,9

I have a mild headache,11,Angioplasty,Yes,8

3. Suggested steps (feel free to modify/choose another approach)

a) EDA – including text (notes - Preprocess review text using spaCy: tokenization, lemmatization, stop-word removal, punctuation removal)

b) Modelling: How will you create a model that can procDecide an approach to process structured data (surgery, number of days etc.) and unstructured data (patient notes). One approach:
Use GPT-2 model as feature extractor (embeddings+attention) and combined with Multi modal attention model to treat text and numeric features
FCNN it can use the other features to learn to predict
Use transfer learning technique to add new layers and train the model on the dataet (new layers weights are trained)
Use cross attention between text and numeric features
Use gated fusion/separate pathways mechanism to deal with feature imbalance
Use dropout layers for robustness and generalization

c) Evaluation

6. Deliverables

1. **Notebook / Python code** with the complete application

2. **Screenshots/outputs** visible in shared colab notebook, without any errors
3. **Presentation deck**: Template provided. Must include business case, solution approach and choices considered, architecture diagram, metric table, explain with diagrams transfer learning process and architecture to fix imbalanced inputs.
4. **README** with setup steps