AI Assistant for Schizophrenia Diagnosis

Simon Thomas
Date: 5/21/25
Period: Yilmaz 1

BACKGROUND

- About 1 in 5 adults in the US live with a mental illness, 57.8 million in 2021 **[4]**
  - Schizophrenia affects approximately 24 million people worldwide **[5]**
- Only 1 in 3 diagnosed with Schizophrenia receive specialist mental health care **[5]**
  - Certain parts of the world have limited mental health services
- Automation of Schizophrenia diagnosis has been studied, not implemented
  - Due to ethical constraints
- However
  - Research has produced Machine learning algorithms that show high accuracy for diagnosis

Issues With AI

- **Automating medical diagnosis poses ethical concerns**
  - Accuracy :
    - How high of an accuracy is good enough
  - Misdiagnosis
    - False negatives can be detrimental
    - False positives can be stressful
- **Loss of Human Oversight**
  - Too much dependence on AI
    - Doctors should have a say
- **Risk of bias raises questions about generalizability of algorithms [2]**
  - Training data fails to take certain demographics into account
    - Different age groups, gender, racial

OTHER SOLUTIONS

Studies	Feature sets	Classifier	Accuracy
Devia et al. (2019)	EEG activity	Linear discriminant analysis	71.00%
Baradits et al. (2020)	Microstate temporal parameters	Machine learning model	82.70%
Siuly et al. (2020)	EMD features	Ensemble bagged tree	89.60%
Akbari et al. (2021)	Phase space dynamic	*K*-nearest neighbor	94.80%
Kim et al. (2021)	Microstate temporal parameters	Support vector machine	75.60%
Lillo et al. (2022)	Microstate and microstate features	Convolutional neural networks	93.00%
Chen X. et al. (2023)	Linear and non-linear measures	Support vector machine	89.00%
Tian Y. et al. (2023)	Structural Brain MRIs	Convolutional neural network	98.70%

OTHER METHODS OF DIAGNOSIS

- **Neural Network Classification [9]**
  - 93% accuracy for diagnosis, using microstate analysis - a microstate is the state of the brain at a very short period of time (80-100 ms)
- **Linear Discriminant Analysis Classification [7]**
  - 71% accuracy for diagnosis, based off of brain EEG activity - EEG readings measure the electric pulses within our brain
- **Convolutional Neural Network Classification [11]**
  - 98% accuracy for diagnosis, using structural brain MRIs

Why IS Mine BETTER?

- Schizophrenia Diagnosis using Convolutional Neural Network (CNN) classifier + Natural Language Processing (NLP)
  - Analysis of brain MRI scans
  - Medical records (data like age, gender, etc.)

Novelty

- Incorporates NLP into schizophrenia diagnosis
  - Schizophrenia shows patterns among different genders and age groups, NLP will take this into account
- “AI Assistant” for diagnosis
  - Tool used by clinicians, not complete automation for diagnosis
  - Identification of subtle patterns which clinicians can take into account

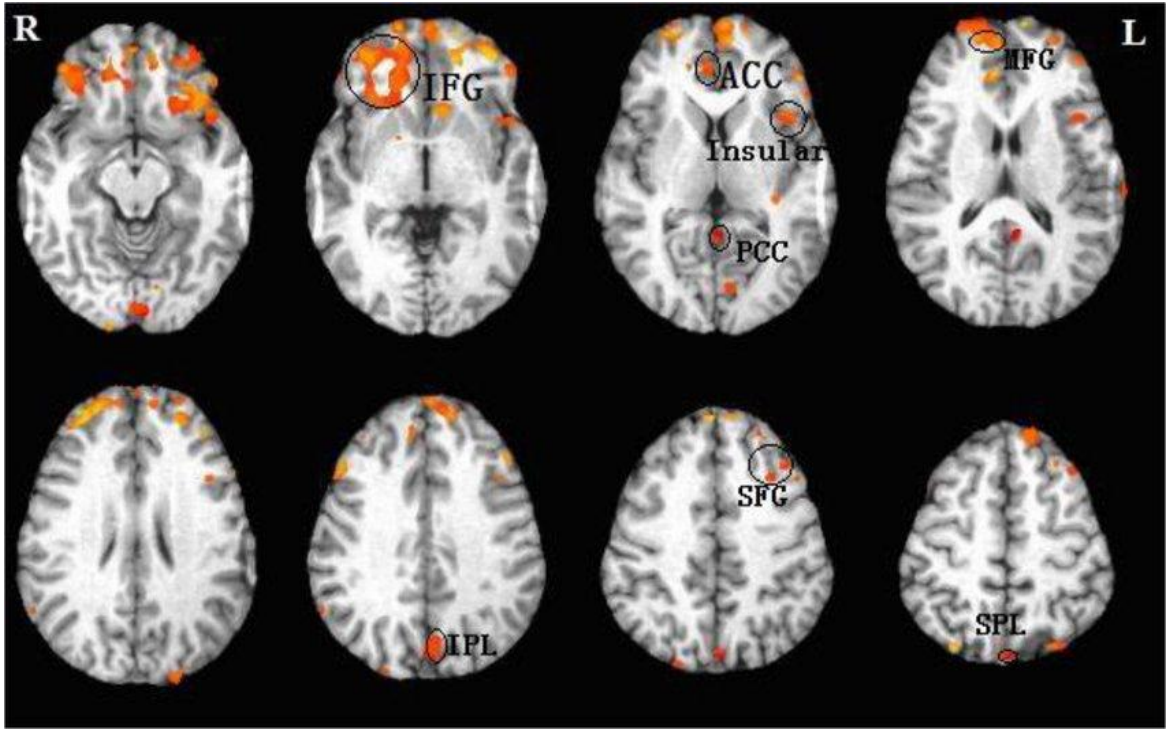
Impact

- **Early detection of Onset Schizophrenia**
  - Faster diagnosis than human diagnosis
- **Higher Confidence in Final Diagnosis**
  - Specialists are reassured in their decisions
- **Accessibility for Diagnosis**
  - Tool can be provided at a cheap price in regions which may have a shortage of healthcare professionals

METHODS: CNN Input Data

Data from COBRE [1]

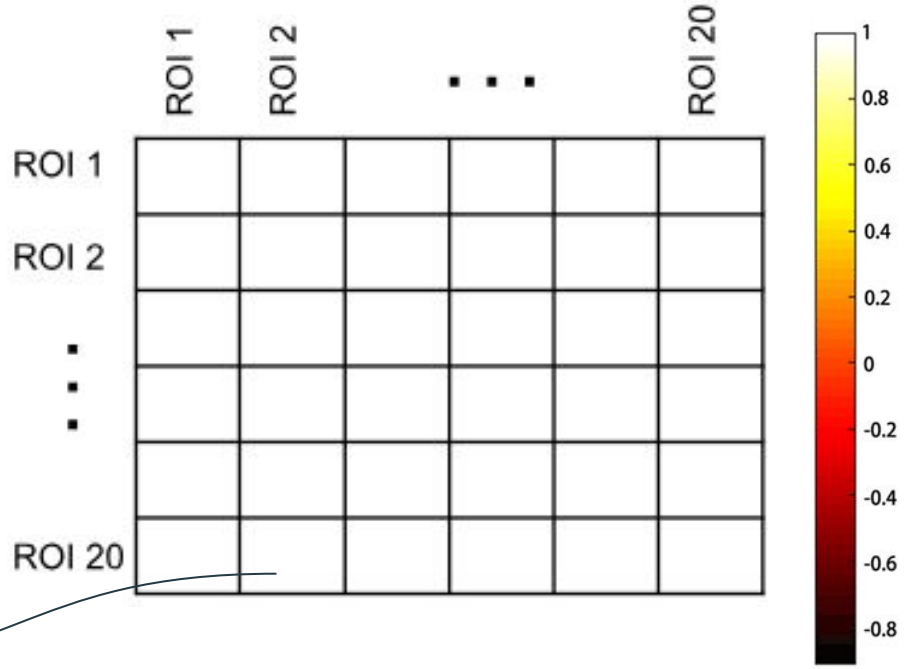
- fMRI imaging results (processed through CNN)
- 20 Regions of Interest



METHODS: CNN Input Data

Data from COBRE [1]

- 20 Regions of Interest
  - 20x20 correlation matrix



Correlation of ROI 2 and ROI 20

METHODS: NLP Input Data

- **Synthetic Data for NLP**
  - Age, Gender, Doctor’s Notes, Family History
- **Positive Example:**
  - "Patient reports auditory hallucinations and paranoid delusions. Difficulty concentrating and social withdrawal observed. Family History: Father has a history of psychosis.”
- **Negative Example:**
  - "Patient describes mild mood swings but no psychotic symptoms. Family history: No known family history."
- **200 Unique Entries**
  - 100 positive, 100 negative
    - 50-50 Train-Test split
    - 50 positive & 50 negative in both the Training Set and Testing Set

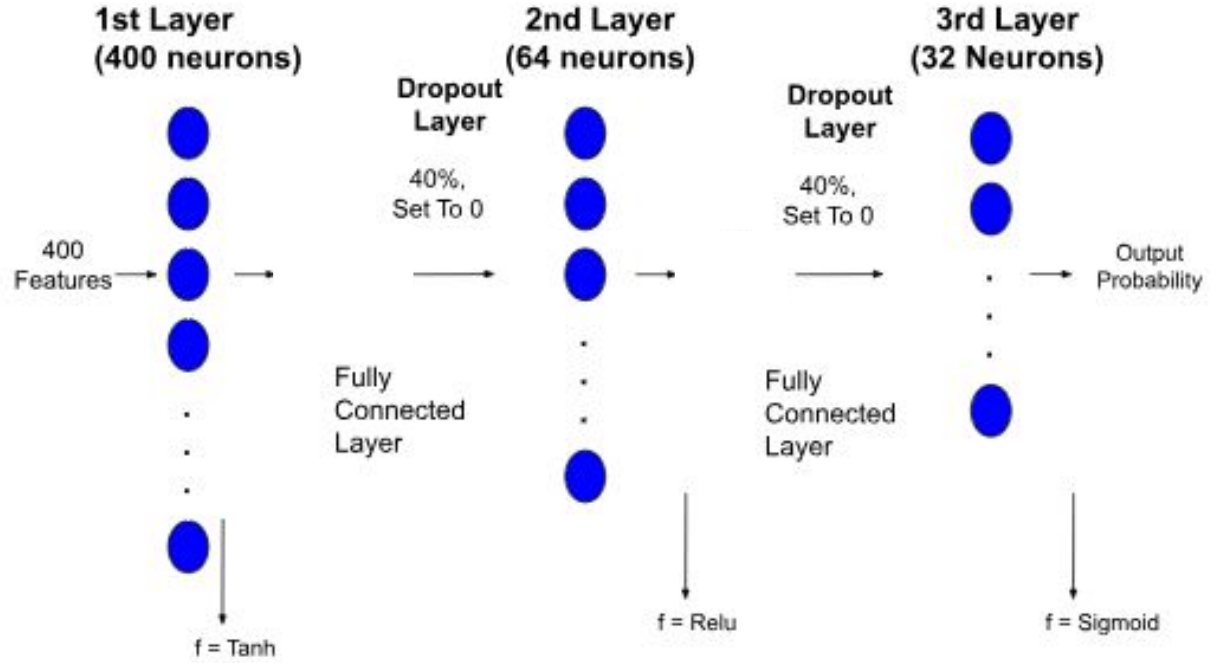
METHODS: General Scheme

- **CNN Evaluation + NLP extraction**
  - Both output probabilities visible to doctors
    - Improves transparency
  - Neural Network training & fine-tuning of NLP algorithm will yield higher accuracy over time
- **Fusion of CNN and NLP through weighted average of probabilities**
  - Doctor can evaluate final probability

METHOD: CNN Architecture

CNN:

Trained using large dataset of labelled “schizophrenic” data consisting of patient MRI brain images, analyzes unlabelled data and outputs probability of SCZ

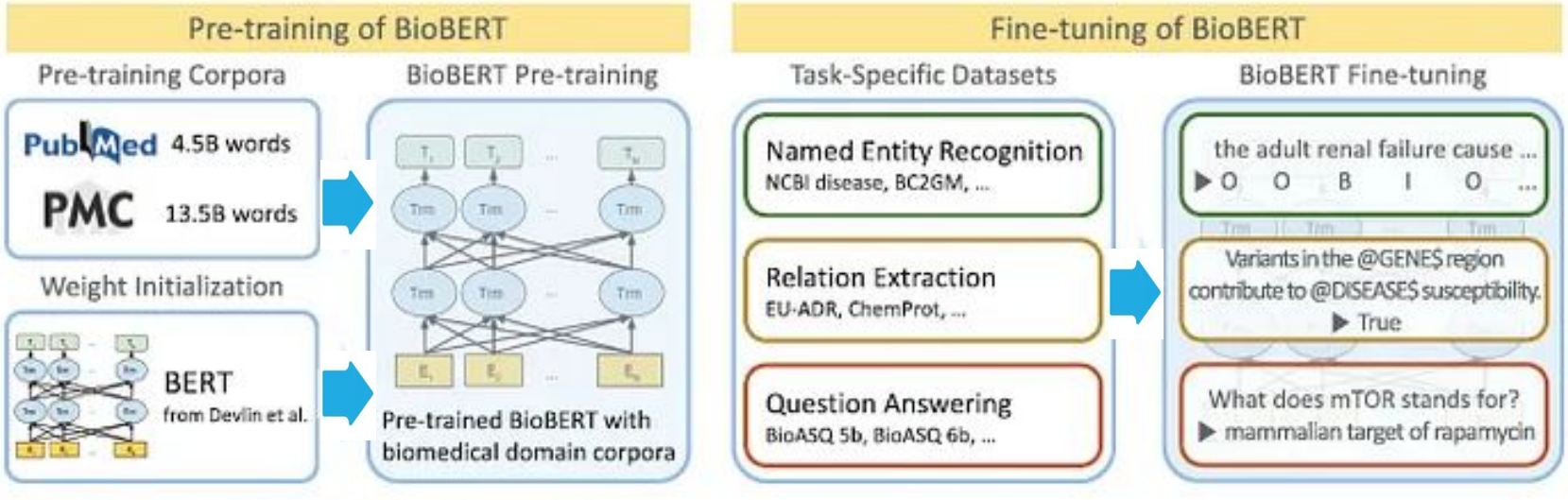


Input = 20x20 Regions of Interest

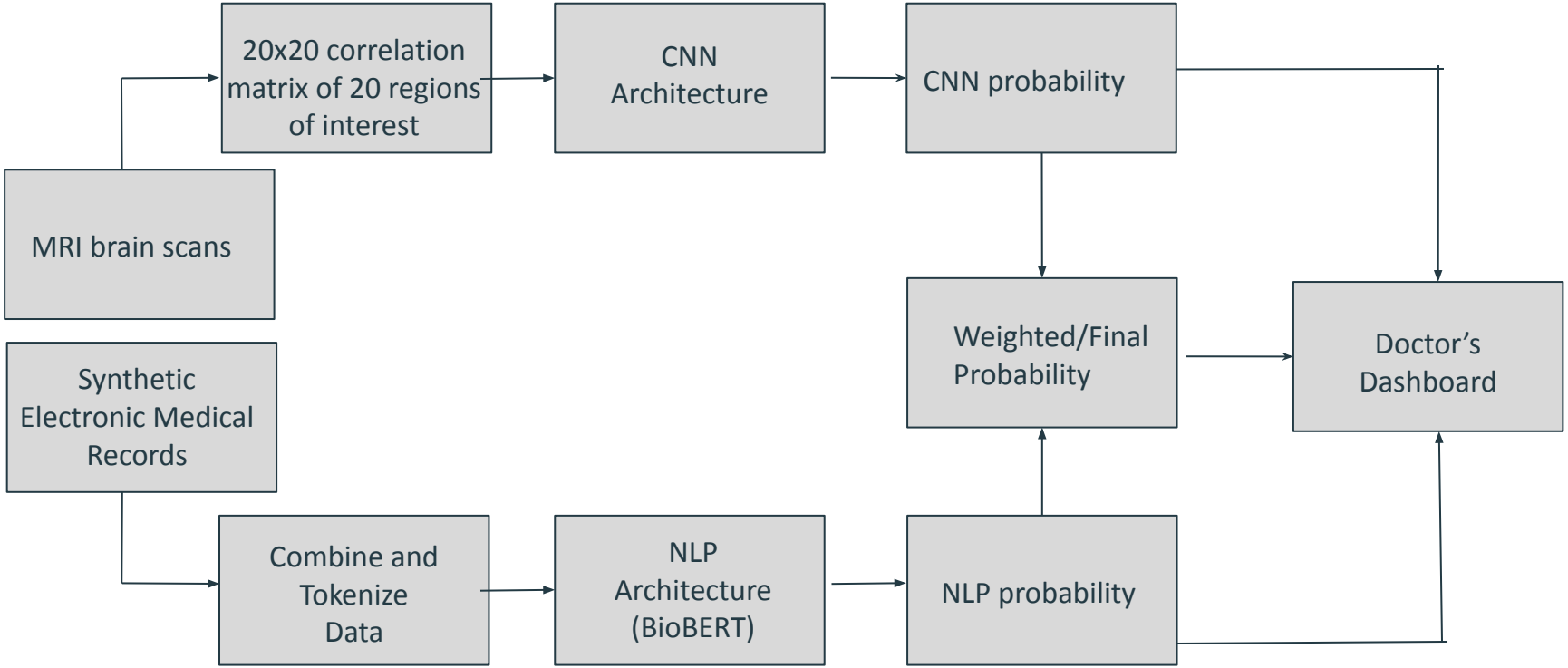
Output of a neuron = f(input*weight + bias)

METHOD: NLP Architecture

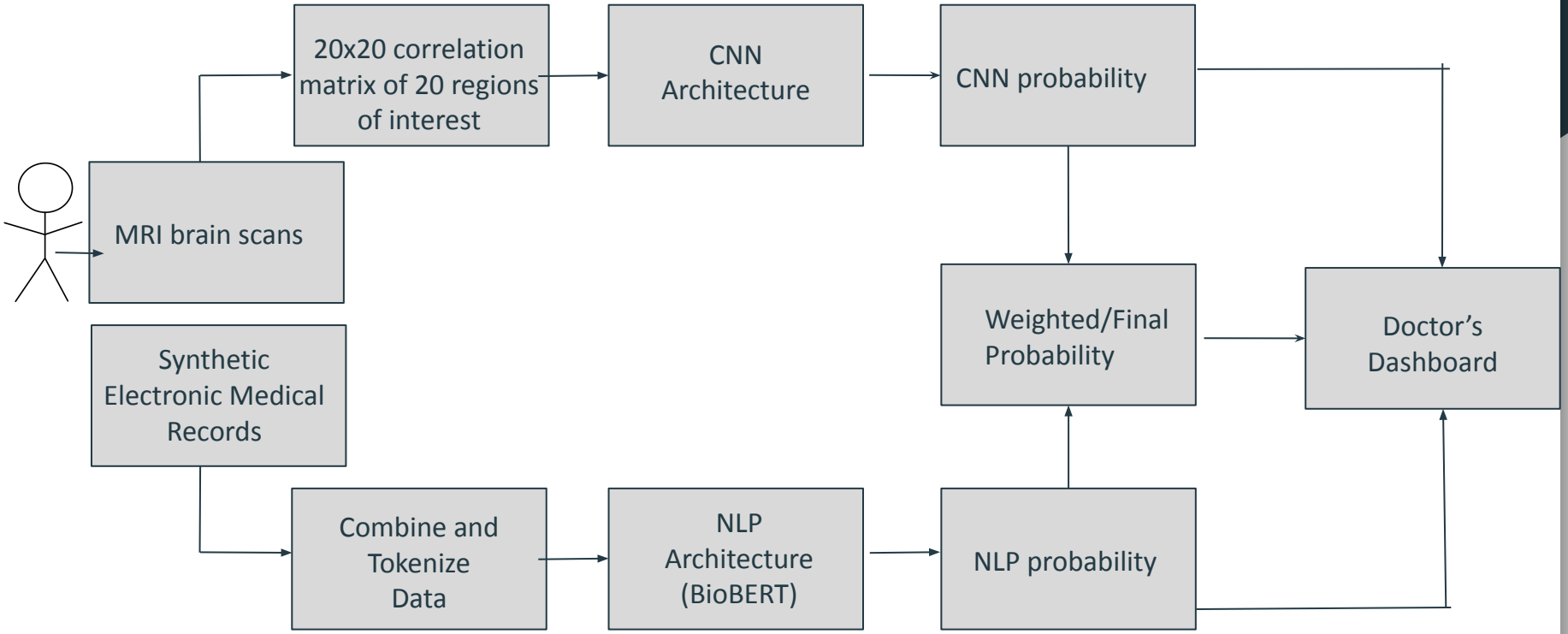
- **NLP (BioBERT)**
  - Concatenation of structured data (age, gender, etc.) and unstructured data (doctor's notes, symptoms, etc.) to output probability vector
- **BioBERT**



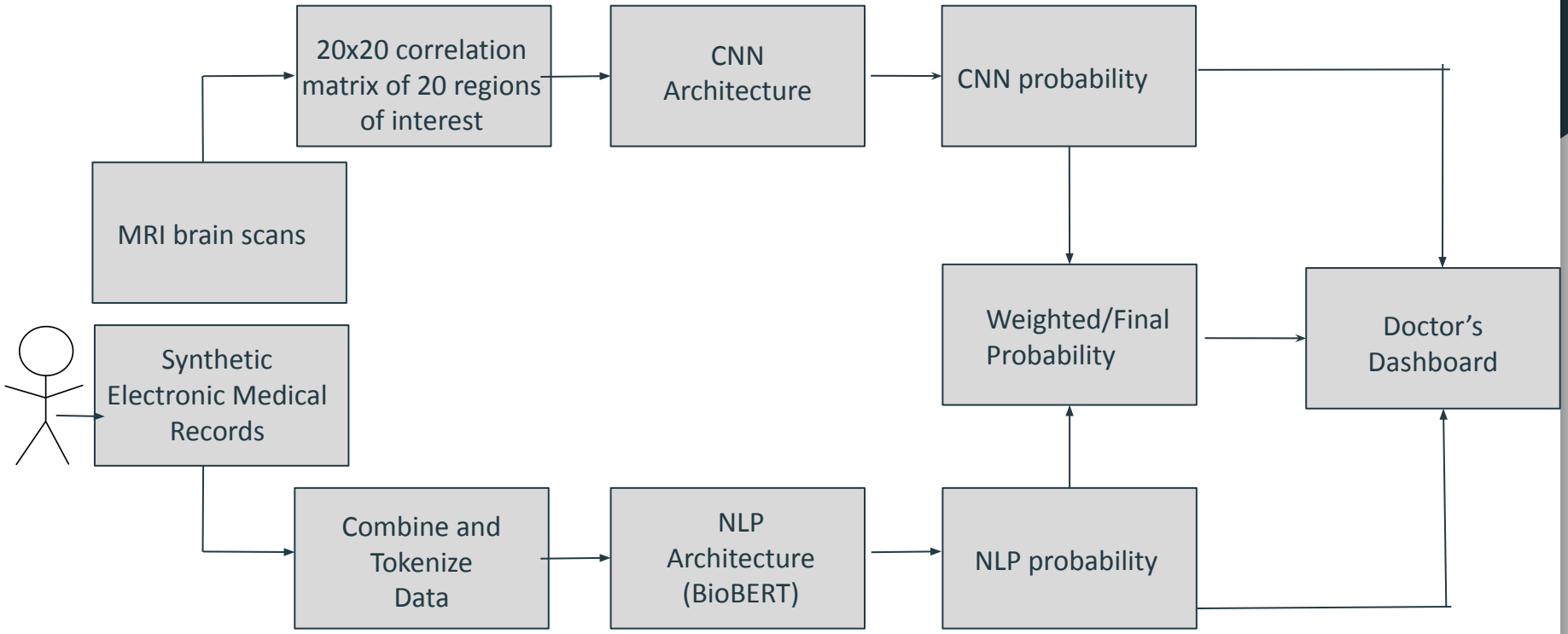
METHOD: Systems Architecture Flowchart



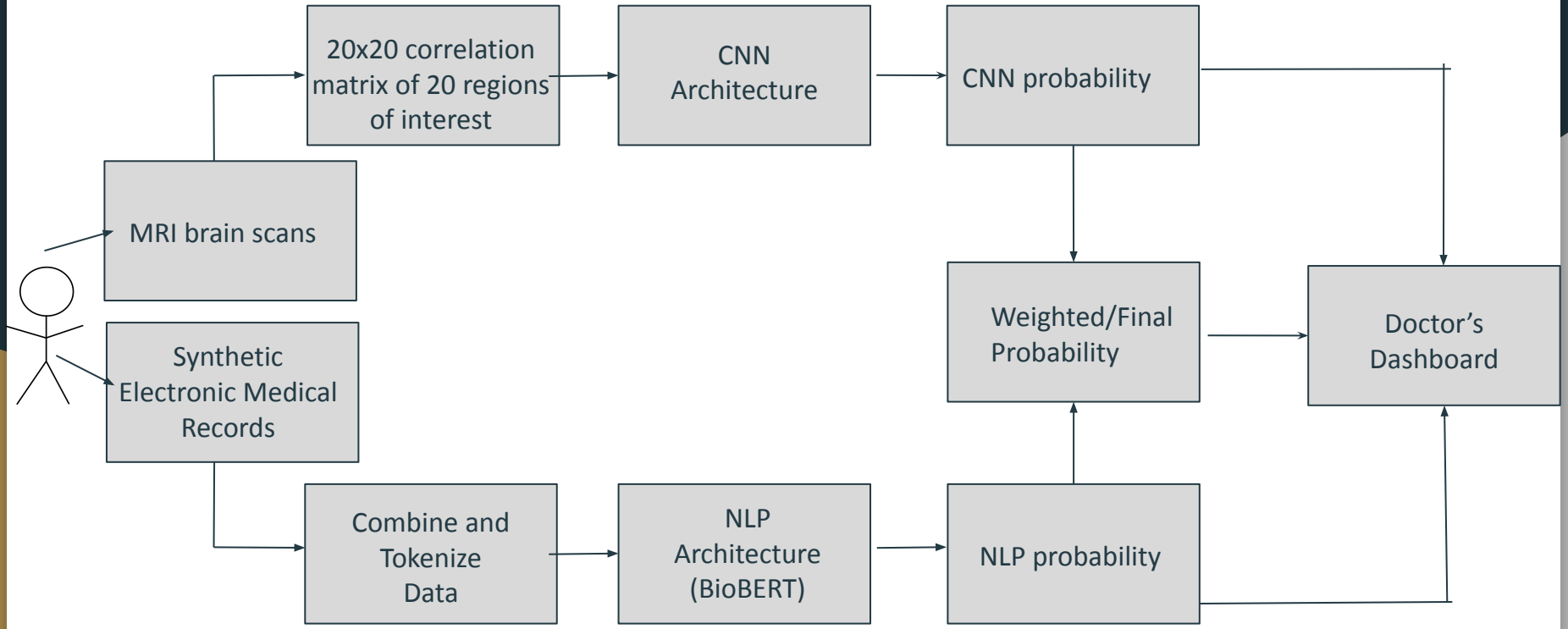
METHOD: Systems Architecture Flowchart



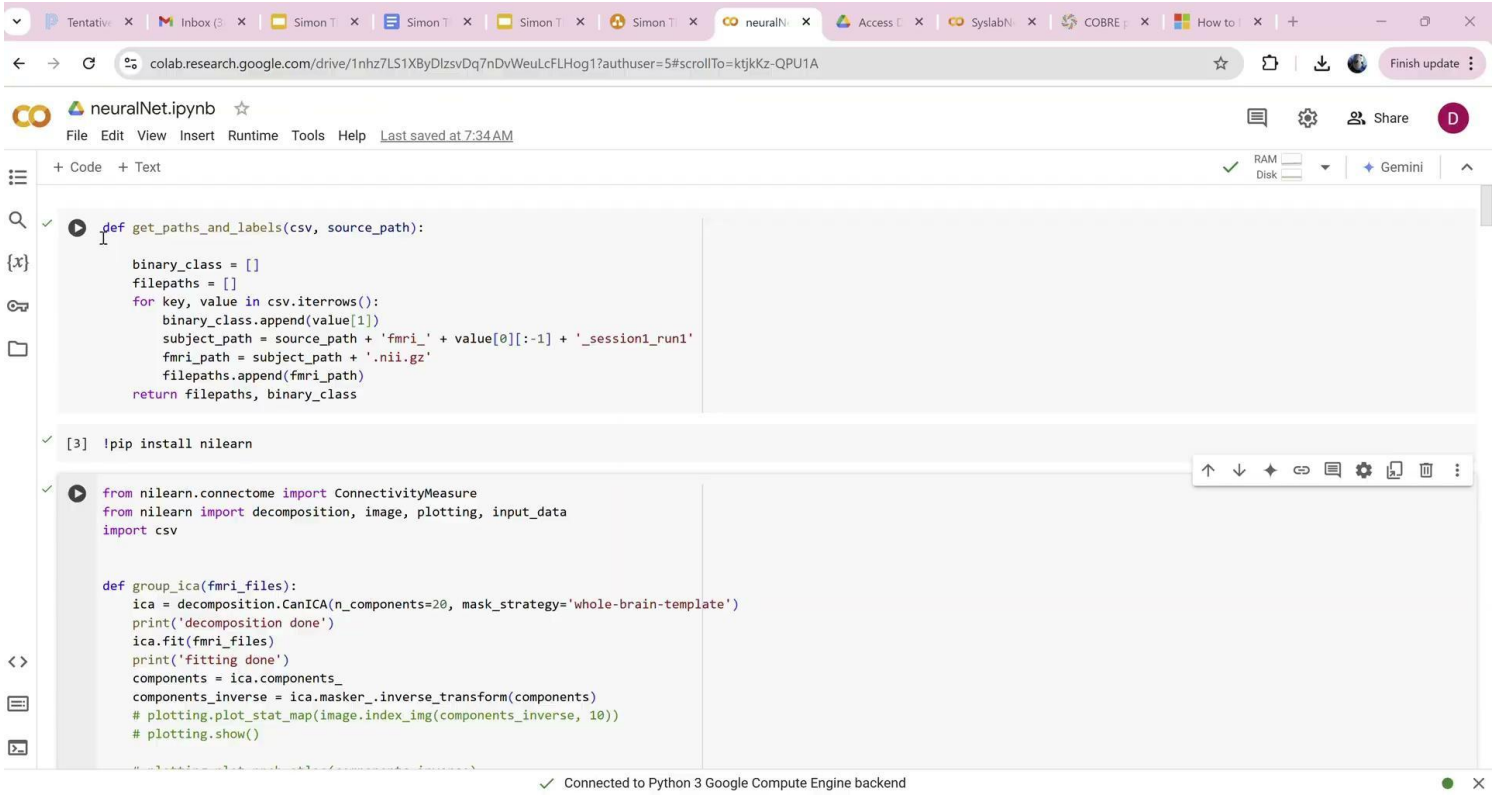
METHOD: Systems Architecture Flowchart



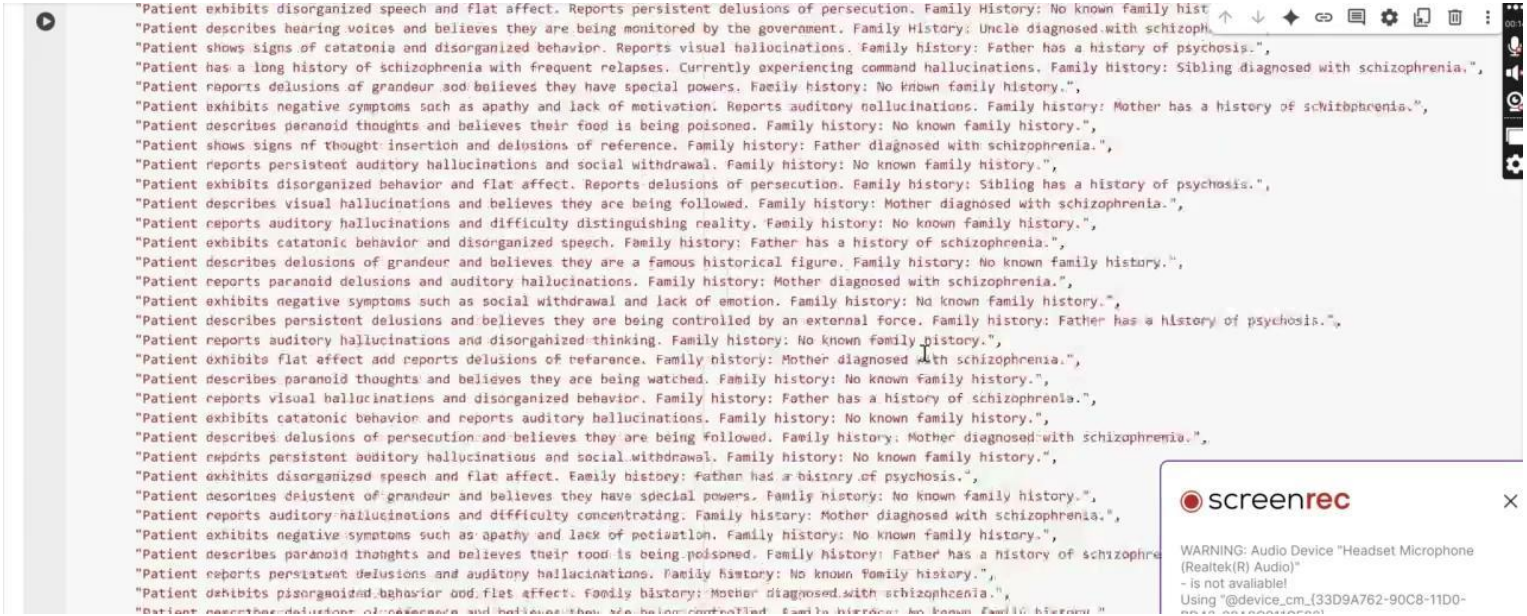
METHOD: Systems Architecture Flowchart



CNN Demo


colab.research.google.com/drive/1nhz7LS1XByDIzsvDq7nDvWeuLcFLHog1?authuser=5#scrollTo=ktjkKz-QPU1A
Finish update
neuralNet.ipynb
File Edit View Insert Runtime Tools Help Last saved at 7:34 AM
Share
+ Code + Text
RAM
Disk
Gemini
def get_paths_and_labels(csv, source_path):
binary_class = []
filepaths = []
for key, value in csv.iterrows():
binary_class.append(value[1])
subject_path = source_path + 'fmri_' + value[0][:-1] + '_session1_run1'
fmri_path = subject_path + '.nii.gz'
filepaths.append(fmri_path)
return filepaths, binary_class
[3] !pip install nilearn
from nilearn.connectome import ConnectivityMeasure
from nilearn import decomposition, image, plotting, input_data
import csv
def group_ica(fmri_files):
ica = decomposition.CanICA(n_components=20, mask_strategy='whole-brain-template')
print('decomposition done')
ica.fit(fmri_files)
print('fitting done')
components = ica.components_
components_inverse = ica.masker_.inverse_transform(components)
# plotting.plot_stat_map(image.index_img(components_inverse, 10))
# plotting.show()
Connected to Python 3 Google Compute Engine backend

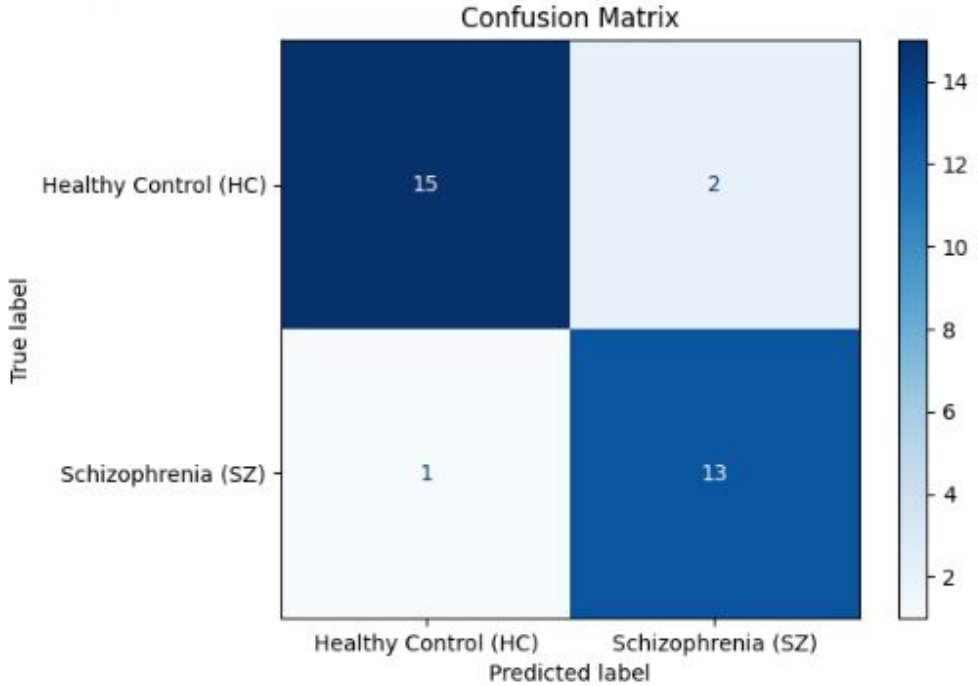
NLP Demo


"Patient exhibits disorganized speech and flat affect. Reports persistent delusions of persecution. Family History: No known family hist
"Patient describes hearing voices and believes they are being monitored by the government. Family History: Uncle diagnosed with schizoph
"Patient shows signs of catatonia and disorganized behavior. Reports visual hallucinations. Family history: Father has a history of psychosis.",
"Patient has a long history of schizophrenia with frequent relapses. Currently experiencing command hallucinations. Family history: Sibling diagnosed with schizophrenia.",
"Patient reports delusions of grandeur and believes they have special powers. Family history: No known family history.",
"Patient exhibits negative symptoms such as apathy and lack of motivation. Reports auditory hallucinations. Family history: Mother has a history of schizophrenia.",
"Patient describes paranoid thoughts and believes their food is being poisoned. Family history: No known family history.",
"Patient shows signs of thought insertion and delusions of reference. Family history: Father diagnosed with schizophrenia.",
"Patient reports persistent auditory hallucinations and social withdrawal. Family history: No known family history.",
"Patient exhibits disorganized behavior and flat affect. Reports delusions of persecution. Family history: Sibling has a history of psychosis.",
"Patient describes visual hallucinations and believes they are being followed. Family history: Mother diagnosed with schizophrenia.",
"Patient reports auditory hallucinations and difficulty distinguishing reality. Family history: No known family history.",
"Patient exhibits catatonic behavior and disorganized speech. Family history: Father has a history of schizophrenia.",
"Patient describes delusions of grandeur and believes they are a famous historical figure. Family history: No known family history.",
"Patient reports paranoid delusions and auditory hallucinations. Family history: Mother diagnosed with schizophrenia.",
"Patient exhibits negative symptoms such as social withdrawal and lack of emotion. Family history: No known family history.",
"Patient describes persistent delusions and believes they are being controlled by an external force. Family history: Father has a history of psychosis.",
"Patient reports auditory hallucinations and disorganized thinking. Family history: No known family history.",
"Patient exhibits flat affect and reports delusions of reference. Family history: Mother diagnosed with schizophrenia.",
"Patient describes paranoid thoughts and believes they are being watched. Family history: No known family history.",
"Patient reports visual hallucinations and disorganized behavior. Family history: Father has a history of schizophrenia.",
"Patient exhibits catatonic behavior and reports auditory hallucinations. Family history: No known family history.",
"Patient describes delusions of persecution and believes they are being followed. Family history: Mother diagnosed with schizophrenia.",
"Patient reports persistent auditory hallucinations and social withdrawal. Family history: No known family history.",
"Patient exhibits disorganized speech and flat affect. Family history: Father has a history of psychosis.",
"Patient describes delusions of grandeur and believes they have special powers. Family history: No known family history.",
"Patient reports auditory hallucinations and difficulty concentrating. Family history: Mother diagnosed with schizophrenia.",
"Patient exhibits negative symptoms such as apathy and lack of motivation. Family history: No known family history.",
"Patient describes paranoid thoughts and believes their food is being poisoned. Family history: Father has a history of schizophre
"Patient reports persistent delusions and auditory hallucinations. Family history: No known family history.",
"Patient exhibits disorganized behavior and flat affect. Family history: Mother diagnosed with schizophrenia.",
screenrec
WARNING: Audio Device "Headset Microphone (Realtek(R) Audio)"
- is not avaliable!
Using "@device_cm_{33D9A762-90C8-11D0-

CNN RESULTS

- CNN Algorithm for Schizophrenia Diagnosis
  - Outputs a final probability which can assist doctor in evaluation
  - 90.3% accuracy

```
Accuracy: 0.9032258064516129
Correct HC: 15
Correct SZ: 13
False HC: 1
False SZ: 2
```



NLP RESULTS

- NLP Algorithm for Schizophrenia Diagnosis
  - Outputs a final probability which can assist doctor in evaluation
  - 97.91% probability
- Reliable for use in Regions of the world with less healthcare
  - Can be trained easily on more realistic dataset
  - A diagnosis machine with high accuracy is better than no evaluation

```
Training the model on synthetic data...
```

```
[57/57 07:51, Epoch 3/3]
```

Epoch	Training Loss	Validation Loss
1	0.675900	0.475052
2	0.322600	0.044813
3	0.074000	0.013892

```
Evaluating the model on test data:
```

```
[19/19 00:27]
{'eval_loss': 0.01389189250767231, 'eval_runtime': 28.6163, '

Predicted Probability of Schizophrenia for the new patient:
0.9791
```

Limitations

- Medical Records like clinician notes are subjective
  - Difficult to obtain medical records
    - Will be used at a hospital setting only
- Synthetic Data
  - Not a true connection between MRI probability and NLP probability
    - Works better using dataset with both MRI and NLP Data for each patient

CONCLUSION AND FUTURE WORK

- **Problem - Automation of schizophrenia diagnosis can be biased and result in loss of human oversight**
  - This project produced a tool for schizophrenia diagnosis based on a diverse data set, and allows for doctor interpretation of probabilities rather than binary output (Positive/Negative)
- **Future Considerations**
  - Dataset with both fMRI and Medical records (patient age, gender, etc.)
  - Mental Illness train and test cases (not Schizophrenia)

References

Dataset:

[1] https://figshare.com/articles/dataset/COBRE_preprocessed_with_NIAK_0_12_4/1160600

References:

[2]https://www.cambridge.org/core/journals/psychological-medicine/article/computing-schizophrenia
[3]https://www.sciencedirect.com/science/article/pii/S0957417422013835
[4]https://www.nimh.nih.gov/health/statistics/mental-illness
[5]https://www.who.int/news-room/fact-sheets/detail/schizophrenia
[6]https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11024310/#B29
[7]]https://ieeexplore.ieee.org/document/8701676
[8]https://www.sciencedirect.com/science/article/pii/S1746809421005140?via%3Dihub
[9]https://www.sciencedirect.com/science/article/pii/S0957417422013835?via%3Dihub
[10]https://www.sciencedirect.com/science/article/pii/S092099641730302X?via%3Dihub

THANKS!
Q&A