Programming for Artificial Intelligence

Company Name:

"CogniVue"

Prepared By:

Name	Roll Number	Email Address
Ayesha Siddiqa	241419	241419@students.au.edu.pk
Dua Mehmood	241391	241391@students.au.edu.pk
Muattar Imam	241451	241451@students.au.edu.pk
Mahnoor	241453	241453@students.au.edu.pk

Submitted To:

Mam

Email: @students.au.edu.pk

1. Introduction

CogniVue is an end-to-end artificial intelligence system designed to predict which brain region will become active *before* it happens and associated cognitive state from multichannel electroencephalography (EEG) signals during cognitive tasks. The system combines:

1. **Deep learning** (transformer-based neural networks) for sequence modeling of EEG time series.

2. **Real-time inference** via a Flask/Python backend API.

3. **Professional visualization** through a React-based frontend dashboard.

The primary goal is to demonstrate that **transformer encoder architectures** can learn to predict the next-second dominant brain activation pattern from current EEG windows

2. Motivation and Significance

Current EEG analysis tools (e.g., EEGLAB, BrainVision Analyzer) are primarily designed for offline, post-hoc review of recorded brain signals. They excel at identifying epileptic events, detecting steady-state responses, or computing spectral features, but they do **not** predict future brain states.

Why predict the next brain activation?

- Attention & workload:
  If we can see that attention is about to drop, systems can give quick help (like alerts or breaks) before performance gets worse.
- Brain-computer interfaces:
  Predicting the next brain state helps BCIs be smoother and more responsive by cleaning up noise and anticipating user actions.
- State changes over time:
  Tracking how the brain moves from focused to drifting to drowsy helps design better personalized learning and therapy methods.

3. Dataset

3.1 Dataset Overview

The project uses the **COG-BCI database** (Zenodo record 6874129), an open-access, multi-session, multi-task EEG dataset designed for passive brain-computer interface research.

Parameter	**Value**
Participants	29 subjects
Sessions per subject	3 sessions (~1 week apart)
Cognitive tasks	4 types: N-back (memory), MATB-II (attention/workload), PVT (vigilance), Flanker (conflict resolution)
EEG sampling rate	512 Hz
Channels	~64 (10–10 extended montage)
File format	BIDS-compatible (.set/.fdt, .edf, or HDF5)
Total duration	>100 hours of raw EEG
Signal characteristics	Continuous, task-driven, with event markers

3.2 Dataset Description

The study dataset consists of a subset of the COG-BCI database, focusing on selected participants (e.g., sub-01, sub-02, sub-03) and all available sessions and tasks for those subjects, stored under a structured `data/raw/` directory following BIDS conventions.

EEG Signals:

- Loaded from BIDS-compatible EEG files (`.edf` or `.set`) with ~64 scalp channels.
- Sampling rate: ~512 Hz.
- Channels mapped to standard 10–20 / 10–10 electrode layout.
- Task events and experimental conditions obtained from event/behavior JSON files and trigger lists.

Cognitive Tasks:

- **N-back (memory):**
  A sequence of items (like letters) is shown, and the subject must decide whether the current item matches the one shown *N* steps earlier (e.g., 2-back = compare with item two positions back). This strongly engages working memory and attention
- **MATB-II (attention/workload):**
  A multitasking cockpit-style simulation where the subject monitors several panels at once (e.g., tracking gauges, responding to alarms, managing controls). It is designed to increase mental workload and divided attention.
- **PVT (vigilance):**
  A simple reaction-time test: the subject stares at a screen and presses a button as soon as a stimulus appears, often for a long, boring session. It measures sustained alertness and sensitivity to drowsiness and fatigue.
- **Flanker (conflict resolution):**
  The subject sees a row of arrows or letters and must respond based on the middle one while ignoring the surrounding "flankers," which can be congruent (same direction) or incongruent (different direction). This probes response conflict and executive control.

3.3 Data Cleaning and Preprocessing

Loading and Filtering

· Raw EEG recordings for each subject and session were loaded using MNE within a dedicated preprocessing module.

· Band-pass filter 1–40 Hz applied to remove slow drifts and high-frequency noise.

· Channels with abnormally high variance automatically marked as bad

· Notch filter at 50/60 Hz to remove mains interference.

Normalization and Storage

· Each window normalized using z-score (per-channel) to reduce scale differences and stabilize training.

· Full preprocessing pipeline executed for all selected subjects and sessions.

· Resulting windows, labels, sampling rates, and metadata saved in `.pt` (PyTorch) or `.npz` format under `data/processed/train|val|test/` for use in training.

4. Methodology:

4.1 Problem Formulation

Supervised Learning Task: Given a 1-second EEG window at time , predict the dominant activation characteristics at time (next second).

Targets (Multi-Task Prediction):

1. **Next dominant channel:** which of 64 electrodes will have maximum band-power in (64 classes).
2. **Next dominant region:** which anatomical region (frontal, fronto-central, central, temporal-left, temporal-right, parietal, occipital) will dominate in (7 classes).
3. **Next dominant frequency band:** which canonical band (delta, theta, alpha, beta, gamma) will have peak power in (5 classes).
4. **Next cognitive state (optional):** Focused, Drifting, Drowsy, or Neutral based on band-power ratios and task context (4 classes).

4.2 Data Acquisition & Raw Signal Processing

4.2.1 Raw Data

INPUT: BIDS files (*.set, *.fdt) from Zenodo COG-BCI
Library: MNE-Python v1.5+
Function: mne.io.read_raw_eeglab() or mne.io.read_raw_edf()
Output: Raw object with shape (64 channels, ~50K-200K samples depending on task duration)
Sampling rate: 512 Hz

4.2.2 Filtering

- Apply a 1–40 Hz bandpass and 60 Hz notch filter to remove drift, muscle noise, and mains interference
- Mark channels with |z-score| > 3 as "bad" (assumed electrode drift, contact loss) and interpolate from neighboring electrodes
- Downsample 512 → 256 Hz with anti-aliasing to cut data size while preserving useful frequencies
- Use an average reference across all channels to reduce common noise

4.3 Feature Engineering

4.3.1 Raw EEG Features

Input window:

- Duration: 1.0 second
- Sampling rate: 256 Hz
- Samples per window: T=256
- Channels: C=64
- **Tensor shape: (T,C)=(256,64)**

- Each window is **1 second of EEG**, 256 samples × 64 channels (256×64256×64)
- Windows use **75% overlap** (hop = 0.25 s)
- Each channel is **z-score normalized** using training-set mean and std to stabilize training and avoid data leakage

4.3.2 Time-Frequency Feature Extraction (Band-Power)

Frequency bands: 5 canonical EEG bands covering 1–45 Hz

Band	Frequency Range	Physiological Meaning
Delta	1–4 Hz	Sleep, deep focus, fatigue
Theta	4–8 Hz	Working memory, attention
Alpha	8–13 Hz	Relaxation, idling, posterior dominance
Beta	13–30 Hz	Active cognition, motor planning
Gamma	30–45 Hz	High-level integration, binding

4.3.3 Brain Region Mapping (Channel → Region)

7 anatomical regions from 10–10 electrode layout:

Region ID	Region Name	Channels	Frontal EEG Locus
0	Frontal	FP1, FP2, Fz, F3, F4, F7, F8	Prefrontal cortex
1	Fronto-central	FC1, FC2, FCz, FC3, FC4, FC5, FC6	Premotor cortex
2	Central	C1, C2, Cz, C3, C4, C5, C6	Primary motor/sensory

3	Temporal-left	T7 (T3), FT7, TP7, P7	Left temporal lobe
4	Temporal-righ t	T8 (T4), FT8, TP8, P8	Right temporal lobe
5	Parietal	P1, P2, Pz, P3, P4, CP1, CP2, CPz, CP3, CP4	Parietal cortex
6	Occipital	O1, O2, Oz	Visual cortex

4.3.6 Label Computation (Target Variables)

For each window , extract features; labels come from window :

Dominant channel :(Channel with highest summed band-power in next window)

Dominant region:(Region corresponding to dominant channel)

Dominant band :(Band with highest global power in next window)

Cognitive state : Uses band-power ratios to assign state:

- **Focused:** and → State ID = 0
- **Drifting:** → State ID = 1
- **Drowsy:** high, low → State ID = 2
- **Neutral:** Otherwise → State ID = 3

4.4 Data Windowing & Train/Val/Test Splitting

4.4.1 Subject-Level Stratified Split

Subject-wise) split to simulate realistic scenarios where models are deployed on new people.

Allocation (29 subjects → 70/15/15):

- **Train:** sub-01 to sub-20 (20 subjects, 69%)
- **Validation:** sub-21 to sub-24 (4 subjects, 14%)
- **Test:** sub-25 to sub-29 (5 subjects, 17%)

For each subject:

- Extract all 3 sessions × 4 tasks = 12 session-task pairs
- All windows from that subject → same split (train/val/test)

4.4.2 Windowing Strategy

Sliding window extraction:

- Window size: 1.0 second (256 samples at 256 Hz)
- Hop size: 0.25 seconds (64 samples) → **75% overlap**
- Total windows per subject:
- Typical session: ~5 min per task → windows/task
- Per subject: windows

Rationale for 75% overlap:

- Captures subtle temporal dynamics; redundancy helps model generalization
- Standard in audio/EEG processing (e.g., spectrograms use 75–90% overlap)

4.4.3 Dataset Statistics

Split	Subjects	Session-Tasks	Est. Windows	Usage
Train	20	240 (20×3×4)	~288K	Model learning; per-channel normalization stats
Val	4	48 (4×3×4)	~58K	Hyperparameter tuning; early stopping criterion
Test	5	60 (5×3×4)	~72K	Final evaluation; never seen during training

Class distribution (Regions):

- Frontal: ~18%
- Fronto-central: ~16%
- Central: ~20%
- Temporal-L: ~8%

· Temporal-R: ~8%

· Parietal: ~18%

· Occipital: ~12%

Imbalance ratio: 2.5× (Occipital least, Central most) → Addressed with weighted cross-entropy loss

4.5 Model Architecture

- We will use a **transformer model with 6 layers** that takes 1-second EEG windows (256 time steps, 64 channels) and learns patterns over time.
- We will join three pieces of information (EEG summary, band-power features, and task ID) into one feature vector and from it predict **channel, region, band, and brain state** at the same time.

4.6 Training

- We will train all four predictions together using one combined loss that mainly focuses on **channel and region accuracy**
- We will use the **AdamW optimizer** with a small learning rate, mini-batches of 64 windows, up to 100 epochs, and early stopping so we stop when validation loss stops improving

4.7 Metrics

- We will measure **accuracy, F1-score, and confusion matrices** for each output head and also check performance per task (N-back, MATB, PVT, Flanker) against random guessing.

5. Expected Results:

Metric	Expected Range	Random Baseline	Expected Interpretation

Region Top‑1 Accuracy	60–75% (expected)	≈14% (7 classes)	Model is expected to reliably predict the next dominant region.
Channel Top‑3 Accuracy	50–65% (expected)	≈4.7% (3/64)	Correct next channel is expected to appear in the top three.
Band Top‑1 Accuracy	55–70% (expected)	20% (5 classes)	Model is expected to capture dominant frequency band well.
State Macro F1	50–65% (expected)	25% (4 classes)	Balanced performance across cognitive states is anticipated.
Overall Outcome	Above all baselines	—	Short‑horizon brain dynamics are expected to be learnable from 1‑s EEG windows.