Single machine learning classifier to detect EEG signals from clinical dataset - Digital Innovation for early diagnosis of epilepsy and potential biomarker for clinical trials

Bartlomiej Chybowski[1,2,3*], Brian Jordan[5], Katie Fisher[5], Iva Peh[1], Alfredo Gonzalez-Sulser[2,4], Javier Escudero[1,3*], Jay Shetty[2,5*]

1. Muir Maxwell Epilepsy Centre, University of Edinburgh; 2. Child Life and Health, CIR, Deanery of Clinical Sciences, School of Medicine, University of Edinburgh; 3. Institute for Imaging, Data and Communications, School of Engineering, University of Edinburgh; 4. Centre for Discovery Brain Sciences, School of Medicine, University of Edinburgh; 5. Royal Hospital for Children and Young People, NHS Lothian; * Joint Supervising authorship

Introduction

- **Infantile Spasms (IS)** are a rare but severe form of epilepsy in babies, characterised by **epileptic spasms** and **abnormal brain activity,** untreated leads to **developmental regression.**
- Early detection and treatment of IS can lead to complete recovery in 20% and better neurodevelopment for 60% of patients . However, it can take **months to receive a diagnosis:**

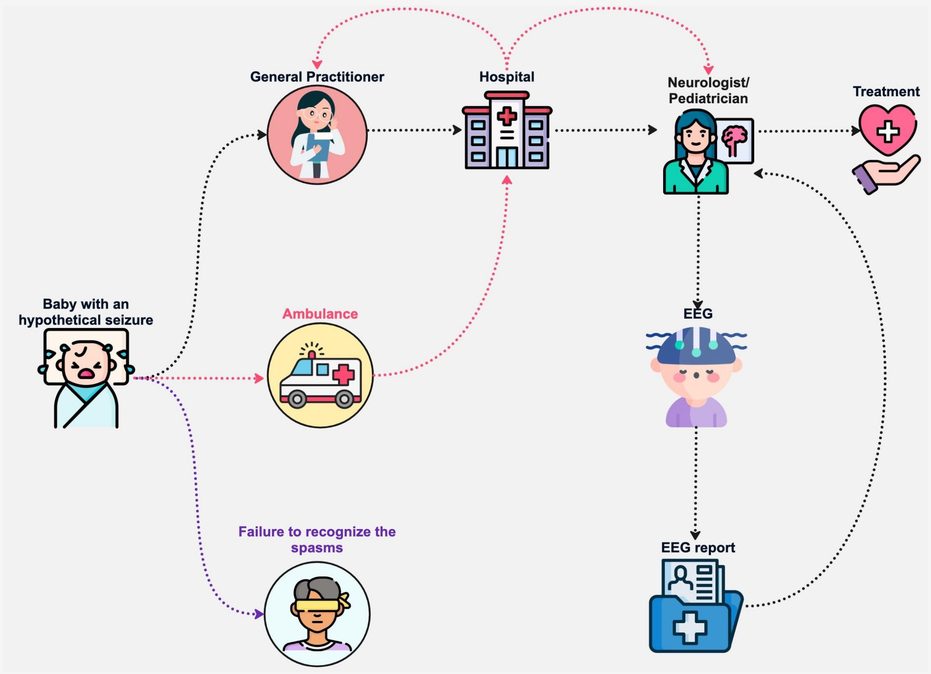

Figure 1. Overview of the current and convoluted path that a family have do to receive diagnosis and treatment of IS.

- Over recent years, statistical analysis and machine learning (ML) techniques advanced significantly, allowing for robust analysis of large **electroencephalogram (EEG)** datasets[1].
- **GOAL: Proof of concept** to determine whether it is possible to use **ML for automatic IS detection**, aiming to assist healthcare professionals in making **faster diagnosis** while **reducing health inequalities.**

Materials & Methods

- Anonymised clinical dataset annotated by neurophysiologists from the Royal Hospital for Children & Young People, Edinburgh:
  - 98 EEG Recordings (50 patients, average 30 mins, total length of 135.5 hours).
  - 42.3 minutes pf seizures (3 second median time).
- An extensive set of 40 unique temporal and frequency features[2,3].
- Eight machine learning models:
  - Seven models individually assessing the recording
  - Majority voting models using predictions of the former.

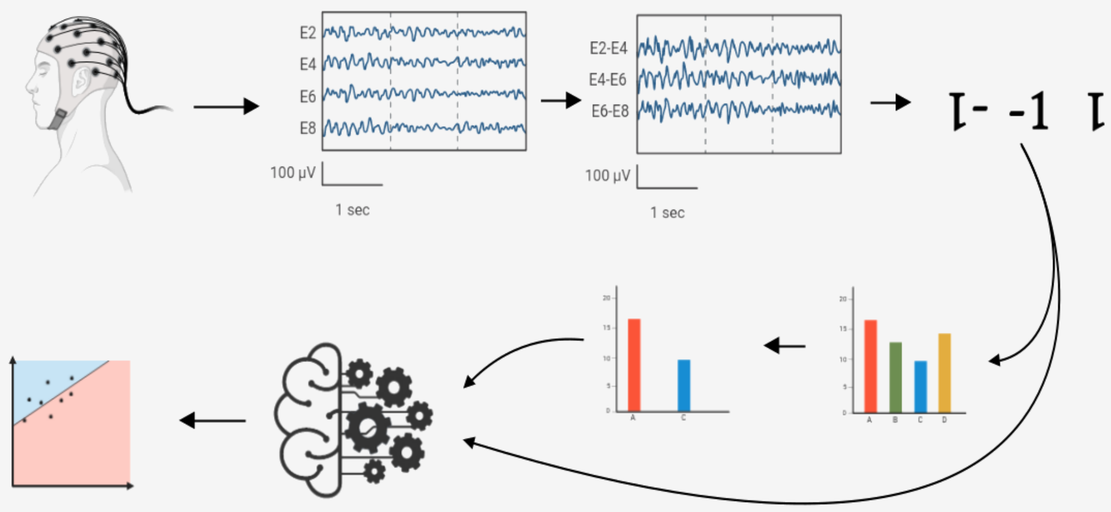

Figure 2. Overview of the research. Steps represent data extraction, pre-processing, feature engineering, feature selection, machine learning training, and prediction.

Results

- We obtained value of **0.76** of receiver operating characteristic curve (ROC AUC) and accuracy of **0.85** for voting model.

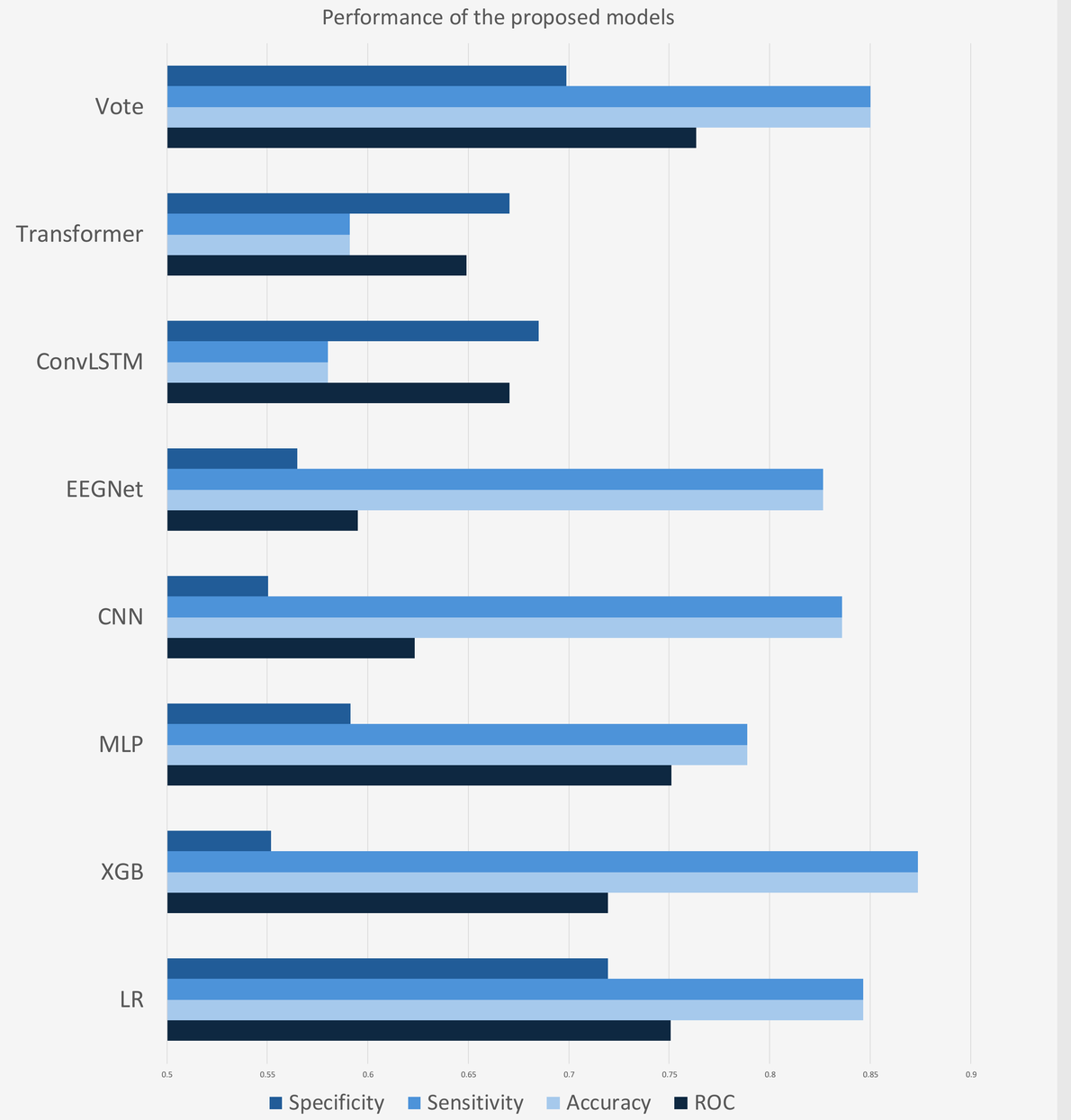

Figure 3. The detection performance of proposed models.

Conclusions

- A majority voting model gave the highest results in terms of ROC AUC and second to the highest in terms of other metrics.
- The results suggest that **a single seizure detection model based on appropriate feature selection and classifiers can achieve stable IS detection performance.**
- This work will be extended through a recently awarded 3-year EPSRC Healthcare Technologies project (EPIC).

References

1. Karoly, P. J. et al. Cycles in epilepsy. Nat Rev Neurol 17, 267–284 (2021)., 2. Anabelle, P. et al., 2020. J Biomed Res 34, 228–239., 3. Hristova, K. et al., 2021. Brain 144, 1576–1589.

Acknowledgements

BC is funded by ERUK-DTC and SIDB.
We also acknowledge a Harmonised Impact Acceleration Award from the University of Edinburgh.

Contact

B.S.Chybowski@sms.ed.ac.uk