David Western

Wallscourt Fellow in Health Technology

Automatic Report-Based Labelling of Clinical EEGs for Classifier Training

D. Western, T. Weber, R. Kandasamy, F. May, S. Taylor, Y. Zhu, L. Canham

3rd November 2021

Problem Statement

EEG is widely used to inform the diagnosis and monitoring of various neurological disorders…

… But is underused due to the dependence on trained human interpreters.

So we want automated EEG analysis for
- Efficiency
- Wider use of EEG, e.g. for screening or 'ruling in' diagnoses

Datasets

Temple University Hospital EEG Corpus (TUEG)

- >30,000 clinical EEG recordings
- Accompanying report text – semi-standardised language
    - Patient history
    - Observations from EEG
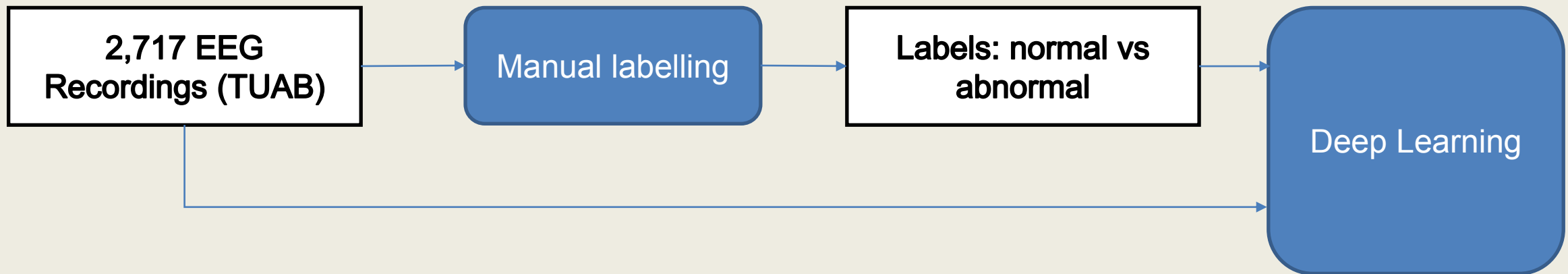   - Impression (opinion on normal vs abnormal)

TUAB

- Curated subset of TUEG for training/evaluating ML classifiers.
- ~3,000 recordings (2,717 for training, 276 for testing)
- Each labelled abnormal/normal based on visual inspection of the single recording.

To make full use of existing clinical data, we don't want to re-inspect and label each recording – use the existing clinical judgement!
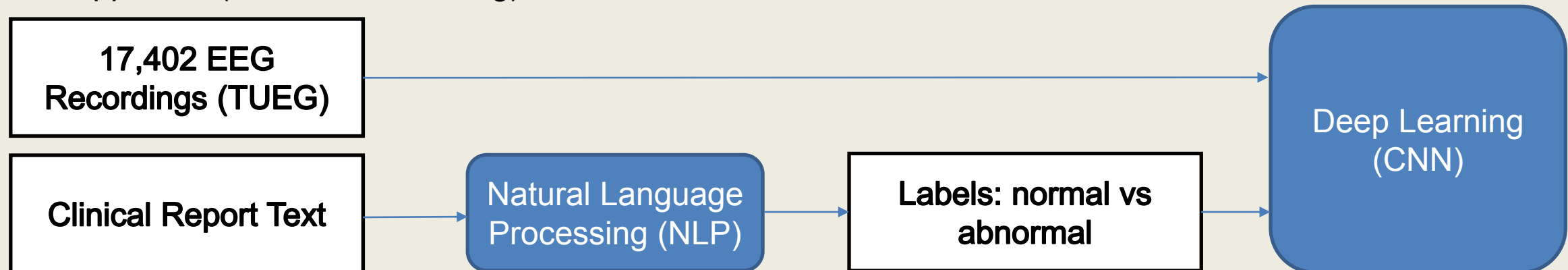
Classifier Training
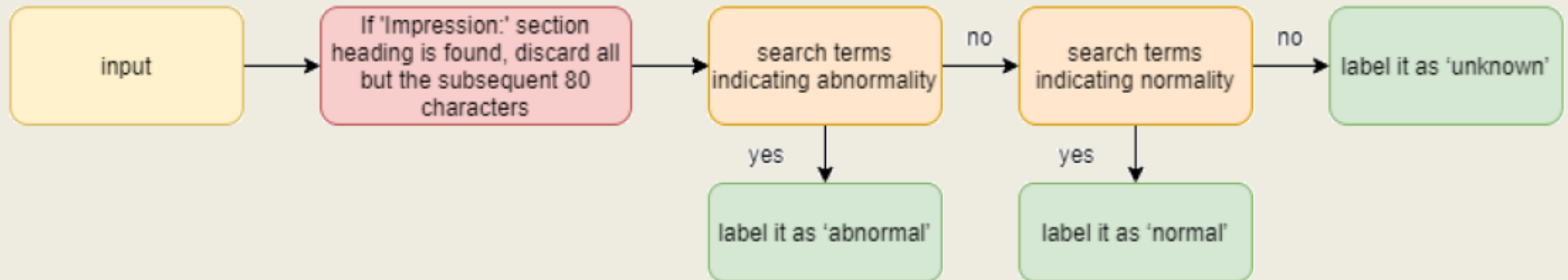
TUAB and conventional approach

```
2,717 EEG Recordings (TUAB) → Manual labelling → Labels: normal vs abnormal → Deep Learning
2,717 EEG Recordings (TUAB) → Deep Learning
```

Our approach (automated labelling)

```
17,402 EEG Recordings (TUEG) → Deep Learning (CNN)
Clinical Report Text → Natural Language Processing (NLP) → Labels: normal vs abnormal → Deep Learning (CNN)
```
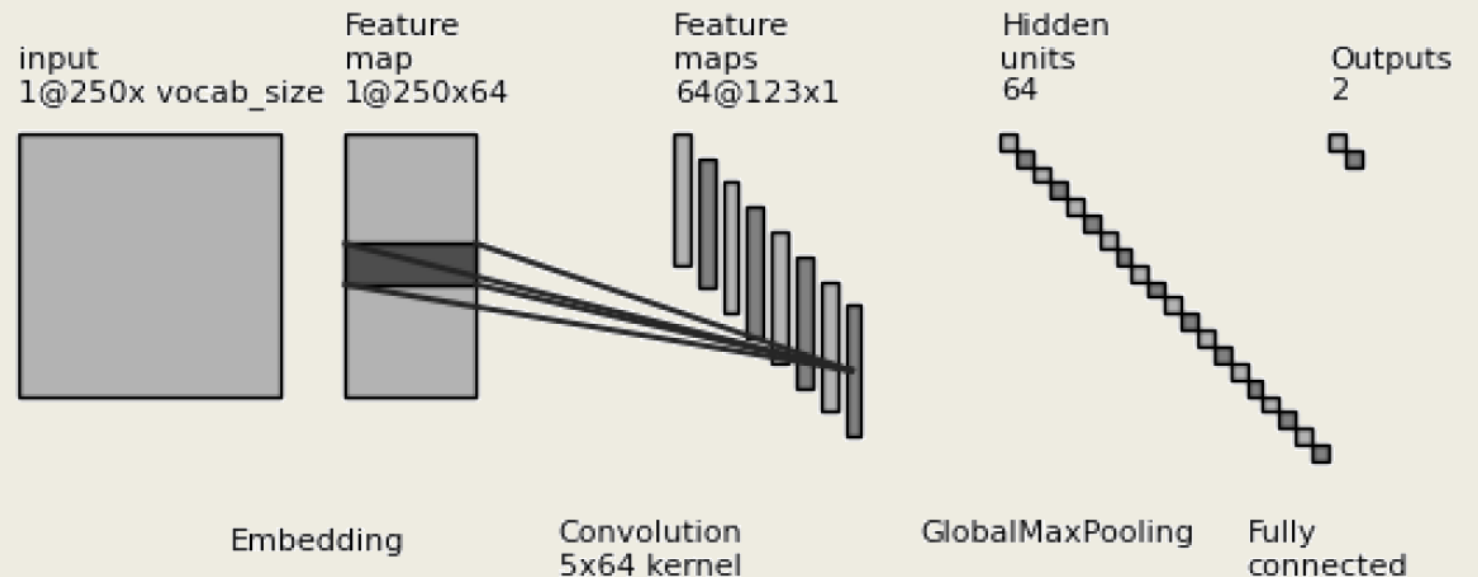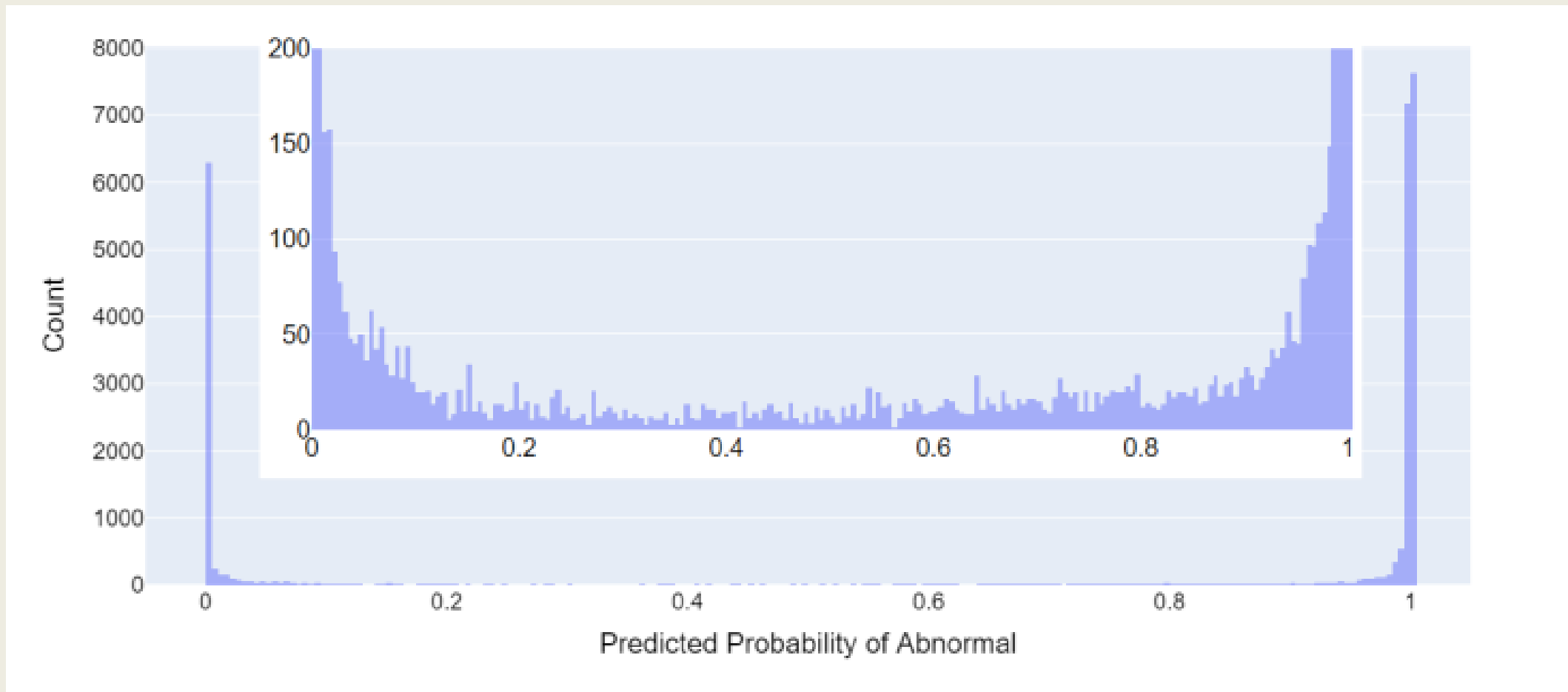
Text Classification

Rule based:

Convolutional Neural Network (NLP):

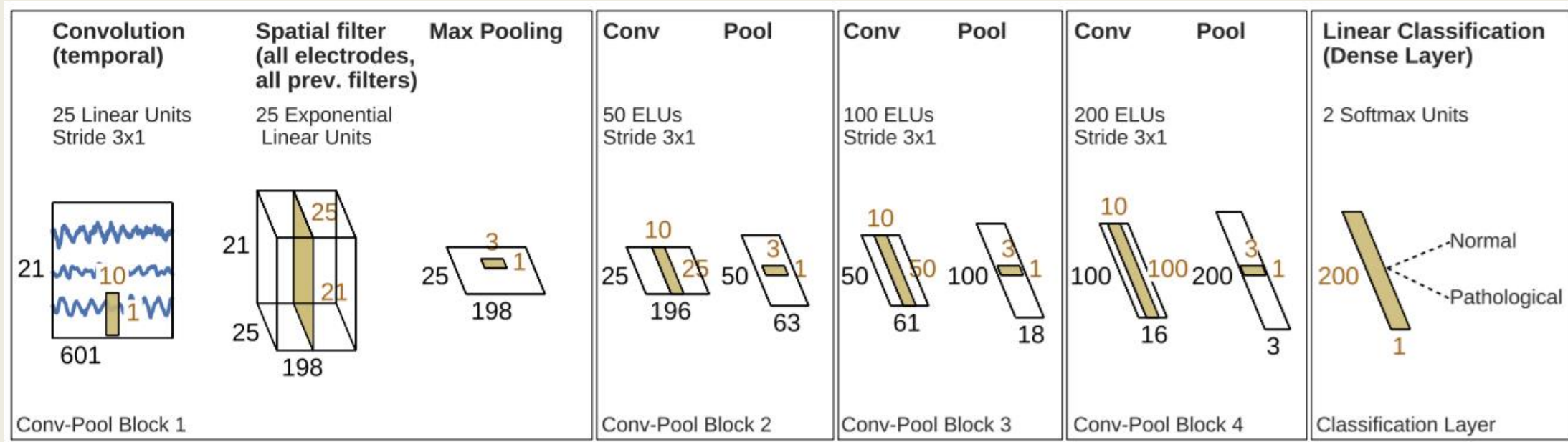# Our NLP labelling 26,387 reports

83 percent are <0.01 (confidently normal) or >0.99 (confidently abnormal).

We discarded the 17 percent in the 0.01-0.99 range.

Impact of Our Approach on EEG Classifier Training

Deep convolutional neural network, as introduced by Schirrmeister et al (2017).

Schirrmeister et al. (2017) Deep Learning with Convolutional Neural Networks for Decoding and Visualization of EEG Pathology. *IEEE SPMB 2017.*

Results

Both tested against TUAB's test
set of 276 manually labelled.

2,717 manually labelled

17,402 auto-labelled

Training Data	Accuracy	Sensitivity	Specificity	AUC
Original TUAB	**81.8 %**	**71.0 %**	90.9 %	0.909
AutoTUAB	77.9 %	53.7 %	**98.3 %**	**0.912**
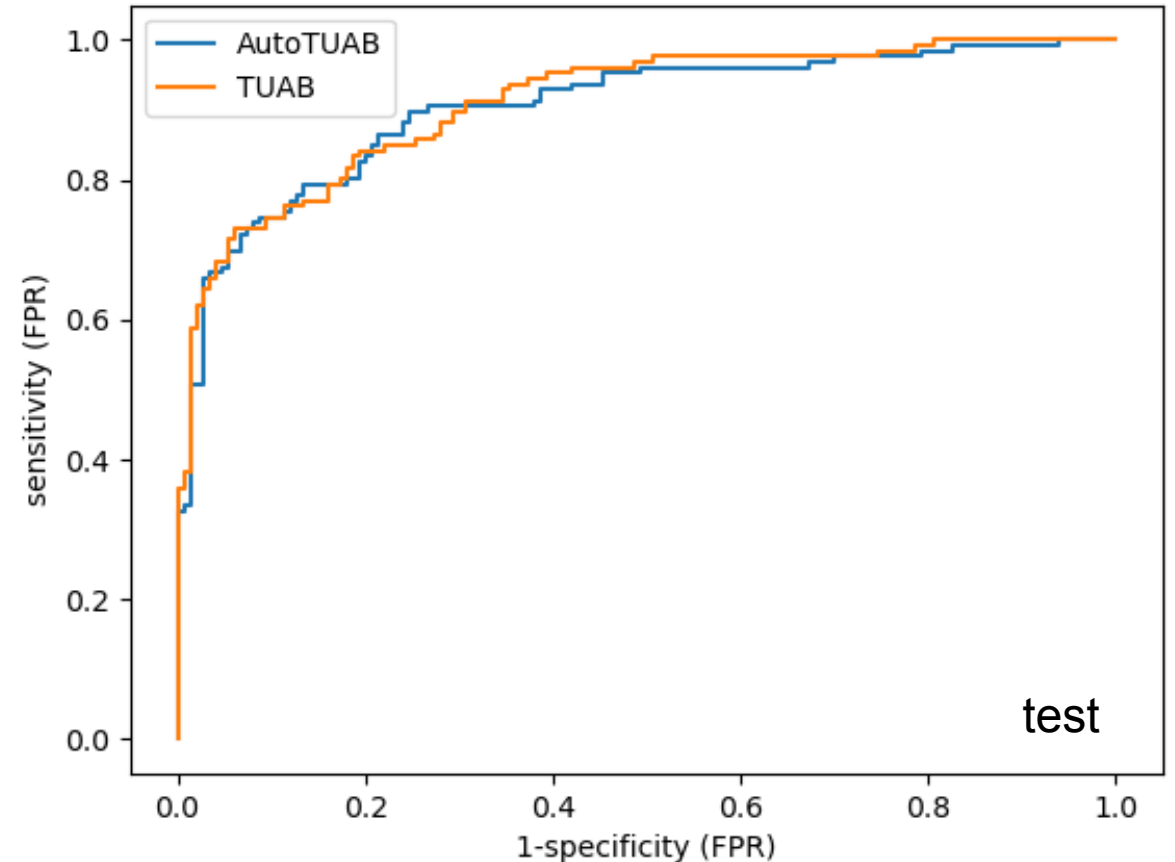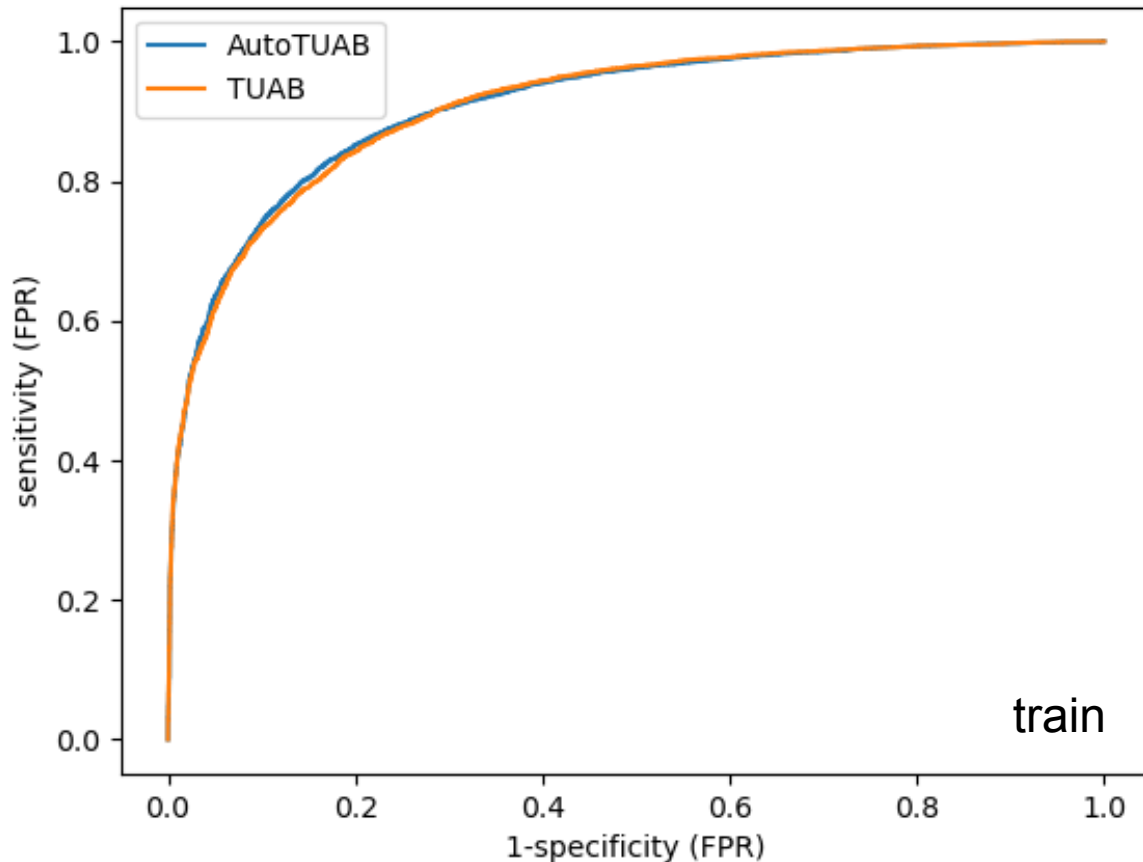
Conclusion

Automated report-based labelling of clinical EEG data

- is feasible.
- yields at least some improvement over a smaller, manually labelled training dataset.

Future Work

Use multiple recordings from within a session to inform the judgement –
more representative of clinical practice.

AutoTUAB test set

github.com/DWonGH/autotuab

Application to other classifier architectures

david.western@uwe.ac.uk

Application to local clinical data