Cambridge Centre for AI in Medicine

UNIVERSITY OF CAMBRIDGE

Real-world clinician use of an Auto-Machine Learning Framework in an NHS trust

NHS Cambridge University Hospitals NHS Foundation Trust

William Pritchard
Cohort 2 Fellow in Clinical Artificial Intelligence & Emergency Medicine Registrar

Background: AutoPrognosis is an AutoML (code-efficient) machine-learning toolkit created by the Cambridge Centre for AI in Medicine.

Project aim: To develop a clinical prediction tool using local data, demonstrating how AutoPrognosis can streamline and simplify the process for clinicians.

Challenges: Limited personal starting skill base- near to no previous coding experience.
Local data access in unmapped and under-construction infrastructure.

Successes:
- Solidification & test-run of local data access infrastructure.
- Upskilling using AutoPrognosis on public dataset.
- Refinement of clinical project suitable for ML prediction & foundation for next steps.

Key reflections:
1. AutoML tools can meaningfully streamline clinicians' use of ML.
2. A non-trivial foundation of data science may still be required (although newer tools further lower this barrier).
3. Some knowledge of what's 'under the hood' is important to spot errors; this could be mitigated with collaboration or robust software.
4. Infrastructure is a critical component to success and efficiency.

Principles of AutoPrognosis

https://github.com/vanderschaarlab/autoprognosis

For classification, regression and survival analysis tasks AutoPrognosis automates:

- Imputation and pre-processing
- Hyperparameter tuning
- Model selection
- Explainability
- Ensemble building
- Demonstrator apps

Alternative toolkits and considerations

PyCaret	FLAML	Auto-ViML
H20 AutoML	EvalML	AutoGluon
TPOT	AutoKeras	MLBox
Auto-sklearn		PyHealth

Tools continue to evolve with more features and reduced coding requirements.

LLMs hugely reduce the barrier to understanding and writing code and code-free toolkits are now available.

Is there an inherent trade-off between accessibility and understanding?

AutoPrognosis allows for creation of an effective predictive ensemble model

This is a classification task for death vs survival in a publicly available COVID dataset

Model code setup

```
study = ClassifierStudy(
    study_name=study_name,
    dataset=training_df,  # pandas DataFrame
    num_iter=2,
    num_study_iter=1,
    imputers=['mean','hyperimpute'] #specify, or remove line for AP to select from all available
    classifiers= ["neural_nets","random_forest","xgboost"] #as above
    timeout=60,
    target="DIED",  # the label column in the dataset
    metric="aucroc",
    random_state=1,
    n_folds_cv=2,
    workspace=workspace,
)
```

Output ensemble architecture

```
study.run()
✓ 9m 53.9s

WeightedEnsemble
WeightedEnsemble(models=[<autoprognosis.plugins.pipeline.ice_nop_scaler_data_cleanup_xgboost object
                         <autoprognosis.plugins.pipeline.ice_nop_scaler_data_cleanup_random_forest
              weights=[0.3636363633057851, 0.6363636357851239])
```

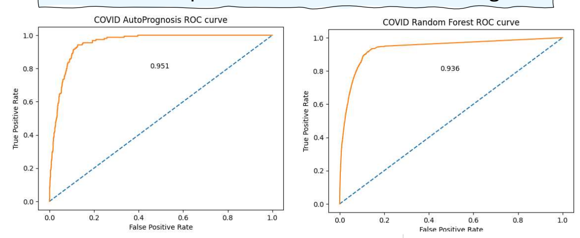
(My) Practice of AutoPrognosis & Reflections

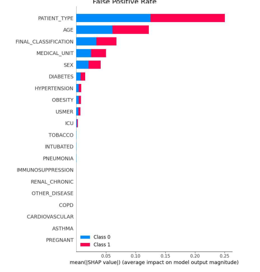
Imputation is effective in comparison with manual inference of missing features e.g. not intubated if sent home, males not pregnant.

Improved performance on single models achievable at a given skill level or time constraint.

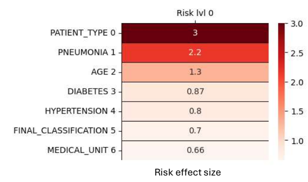
Explainability is not simply for confidence, but allows for rapid identification and correction of errors.

Results in comparison to Scikit-learn modeling





Explainability plugins



Reflections:
There isn't an advertised pre-requisite coding level for AutoPrognosis: I feel I hit a meaningful threshold for understanding when I had learnt to model within Scikit-learn.

While the input parameters for the tool are reasonably simple, some technical knowledge is still required:
(1) For pre-processing (cleaning, feature engineering). This could remain burdensome for complex or unclean datasets.
(2) To configure an environment to interact with the tool.

Better models can be created for a given user's data science skill: a user with moderate skills can create more complex models and a highly skilled user can rapidly test and optimise multiple complex models.

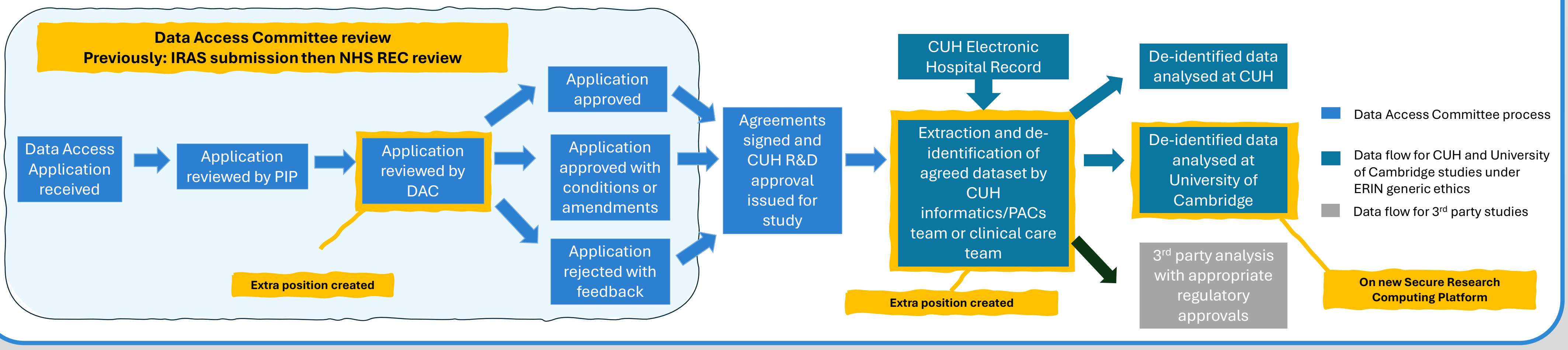
Simple models can be run on a personal PC but complex ensemble models on large datasets benefit from institutional compute power.

New Cambridge University Hospitals data extraction infrastructure

Local Data Access Committee performs the review of studies under a generic ethics framework, which otherwise would have required submission through IRAS to an NHS Research Ethics Committee.

Funding for new positions has expanded the capacity for project review and data extraction.

New Secure Research Computing Platform has been set up to host de-identified data within University of Cambridge (ISO 27001 compliant Secure Research Environment).



Ongoing research

Separately funded 2023 Cambridge Clinical Research Fellowship. Generously supported by The Evelyn Trust and Cambridge University Hospitals Emergency Department Research Fund

Local EHR dataset (10 years, ~800 patients, extensive features) incoming for retrospective study: ML prediction (pre-test probability; using AutoPrognosis) of infective endocarditis in Staphylococcus aureus bacteraemia for rationalisation of echocardiography.

Prospective recruitment of SAB patients for EHR data collection as above **plus** blood sample salvage. Novel molecular pathogen diagnostics adding new features to improve and diversify ML clinical predictions e.g. pre-test probability of IE, LOS, mortality, ITU requirement, early identification of treatment efficacy/inefficacy.