INTRODUCTION

Reviewer: HA

In December 2019, COVID-19, a newly emerging human virus, was discovered in China. Since then, it has spread rapidly across the globe. China's city of Wuhan is believed to be the epicentre of a pandemic that has caused massive hospitalisations (Knight et al. 2020). Besides fever and coughing, the new virus can also cause shortness of breath, pneumonia, as well as various serious respiratory infections (Knight et al. 2020). The growing number of daily contamination cases and massive hospitalizations is a serious threat to already overburdened health systems around the world. Thus, the health system cannot meet the needs of many patients with urgent medical conditions due to overcrowded hospitals MISSING REF: Bhapkar2020. Most of the earlier Covid-19 hospitalization prediction studies focused on predictors such as new daily cases, vaccination rates, pre-existing health conditions, population size and socio-demographic factors to set up hospitalization patterns and trends MISSING REF: Zhou2018. It is essential to monitor and forecast new admissions for inpatients to manage hospital resources efficiently, reduce overcrowding, and improve the quality of care received MISSING REF .... Therefore, in this study, we aim to find the range of predictors that may affect SARS-CoV-2 hospital admissions per NHS Trust as well as to predict Covid-19 new admission inpatients using time series models. Addressing these two research questions will improve NHS performance and patient outcomes by supplying more efficient and higher-quality patient care and optimising the allocation of limited resources to meet the growing demand for hospital places MISSING REF: Huang2019.

Hence, the project focused on the development of overarching task is the development of a model that predicts a NHS Trust's expected new admissions, per day, This led to three core challenges

The models forecast the expected new coronavirus disease admissions per day for the next 15 days.

METHODS

The schematic illustration of *figure ...* outlines the project's data engineering, modelling, and evaluation steps, which underlie the project's research strategy. This section briefly discusses the research strategy, and the steps.

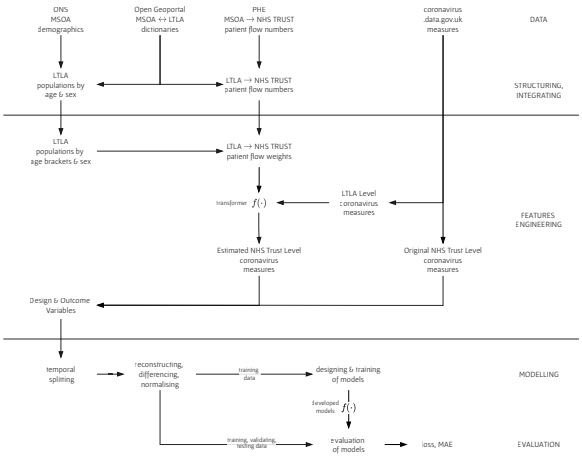

Figure 1: The project's processing, analysis, modelling, and evaluation steps. Please refer to the methodologies section for a brief description of (a) the patient flow weights, and (b) the estimation of NHS trust level measures via flow weights and LTLA level measures. MSOA: middle layer super output area, LTLA: lower tier local authority, ONS: office for national statistics, NHS: national health service, PHE: Public Health England.

RESEARCH STRATEGY

In progress ... includes research strategy (Oates 2006)

DATA ENGINEERING

Data Collection.

The data sources are: (a) the coronavirus.data.gov.uk application programming interface (API) for England's SARS-CoV-2 infections measures, (b) the office for national statistics (ONS) for population estimates, (c) Public Health England (PHE) for the annual intake of patients from one or more middle layer super output areas to an NHS Trust, and (d) the Open Geography Portal (geoportal) for the middle layer super output area (MSOA) ↔ lower tier local authority (LTLA) geographic codes mappings.

Structuring & Integrating.

In progress ...

Features Engineering.

In progress ... The prospective modelling measures exists at either NHS Trust or Lower Tier Local Authority (LTLA) level, but our interest lies is in NHS Trust level admissions, therefore ... the NHS Trust level estimates ...

MODELLING

- Pre-modelling procedures. [This point refers to the pre-ML/pre-forecasting pre-processing. It is separate from the initial pre-processing & features engineering that leads to the raw design matrix and dependent variable.]
- Data mining techniques?
- Statistical techniques?

H.A., J.R., C.O. were discussing:

- Data preparation ... temporal splitting, [rec, diff, ...]
- model design & training ...
  - CNN, LSTM, GRU ... why? briefly discuss type of problem
  - Window ... why? ... forecasting 15 days into the future based on unknown "effective days of history" ... windows allow us to investigate via varying history window sizes

EVALUATION

A brief description of the evaluation metrics. [Results discusses the evaluations in detail]

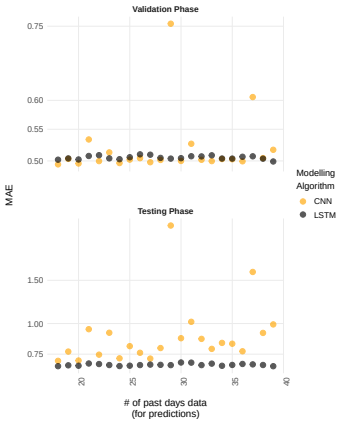

Figure 2: MAE w.r.t. the ...

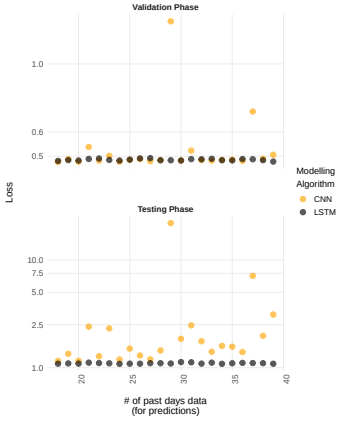

Figure 3: Loss w.r.t. the ...

RESULTS

MODEL EVALUATION

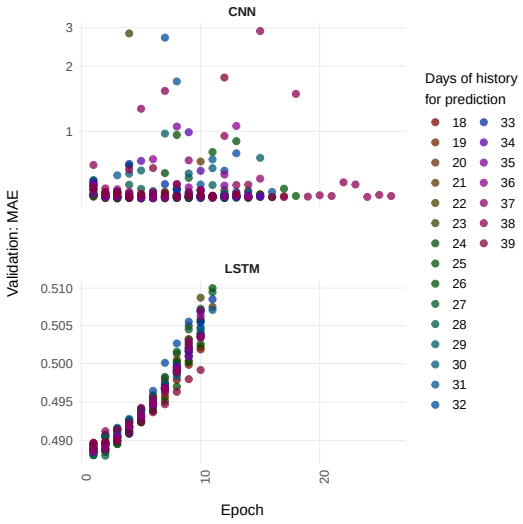

Figure 4: Pending

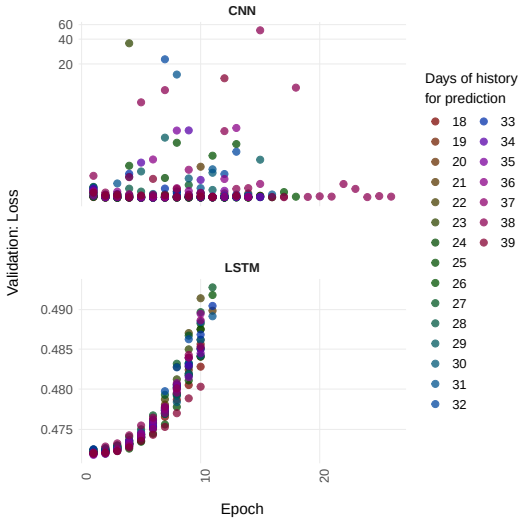

Figure 5: The loss errors

Project guide:

- Explain your results and what your analysis revealed ()
- What implications would your analysis' results have? How do your findings relate to the original question?

BIASES & VALIDITY

- Were there potential biases in your work?
- Validity (remember to discuss what would have been done differently to address identified limitations)

CONCLUSIONS

Project Guide:

- Reflection on the approach taken. (Appropriate?)
- How would you have improved the approach in future? (Alternative methodologies, models, etc)

REFERENCES

Knight, Stephen R, Antonia Ho, Riinu Pius, Iain Buchan, Gail Carson, Thomas M Drake, Jake Dunning, et al. 2020. "Risk Stratification of Patients Admitted to Hospital with Covid-19 Using the ISARIC WHO Clinical Characterisation Protocol: Development and Validation of the 4c Mortality Score." Edited by J Kenneth Baillie, Malcolm G Semple, Peter JM Openshaw, Gail Carson, Beatrice Alex, Benjamin Bach, Wendy S Barclay, et al. *BMJ* 370. https://doi.org/10.1136/bmj.m3339.

Oates, Briony J. 2006. *Researching Information Systems and Computing*. SAGE.