

Minutes - Thursday 27/05 - Pierre Lafaye de Micheaux

Lecturer : Pierre Lafaye de Micheaux

Date : Thursday 27/05/21, 7-8pm

Attendees : Pierre Lafaye de Micheaux, James Cleaver, Md Ruhul Amin Sarker, Peter Morian, Jittinun (Nee) Trairattanasirikul

Minutes

- Walk through our baseline model results
 - RMSE (train data) vs RMSEP - Root Mean Squared Error of Prediction (test data)
 - Features selection - to calculate VIF then decide column with large VIF value to be dropped then re-calculate again
- Video and presentation ideas
 - Audience - lecturers, group of fellow students, senior executives, technical colleagues
 - Not too technical as report already contains technical info, not too much in details of model, potential future employer, peer marking by other groups (target what other teams can learn from us)
 - Agenda: Defining what's the problem, how to resolve, what are the findings at the end
- Can you please share a video from a past term?
- Can you please provide more feedback on our GitHub structure?
 - Put links in each tool
- Black box model - need to be able to explain if use
- Should we normalise our data? Why and why not?
 - Normalisation - 0-1 ranges
 - Standardisation - divided by SD, usually, need to SD
 - We should standardise data if measure in the same unit it will destroy information
 - PV for multiple locations, need to multiply total, good for assumption
- Can you please advise who will be marking?
 - All 3 lecturers
 - Answer questions/assumptions
 - This model is good because of this, know how to use the model
 - Explain how to use graph
- Confidentiality
 - Ideas should not be used outside academic. Write in the report.

Next Actions:

- Grant Colab access to lecturers - James
- Add tool links to Github - Nee
- Add a confidentiality note in the report - Peter
 - The methods, findings and recommendations in this report, as well as the source code used to produce the results in this report, are only to be used for academic purposes. Unless express prior written consent is provided by all authors involved, no part of this report and the associated source code is to be used in the commercial setting.
- Normalise data (not standardise) PV data - Ruhul
- Drop last one-hot columns (last...30min time, season, workday, day, weekend) & VIF & drop columns before regression - Nee

Next meeting: Friday 28/05 with Oliver Nunn (Sponsor)