**Midterm Report**

[**ACM template**](https://www.overleaf.com/latex/templates/acm-conference-proceedings-master-template/pnrfvrrdbfwt)

**(Account: thomassun9917@gmail.com**

**Pwd: cs145fall2020**

**)**

1. Data processing and transformation (jason)

* Fill nas (Check notebook, just use 0s, give justification)
* Check by state (need plot)
* Filter out features (some are of a constant multiple of another, those by rate)
* Smoothing (Tianchen help)
* Polynomial features (sklearn poly, deg 3)

2.Designed & tested models / methods

* Linear (Tianchen
* Poly (Tianchen
* LSTM (LWX)
* Exp smoothing (Thomas)
* SEIR (PX

3.conclusions and findings

-Finding, linear (using tight window for better performance) (Thomas

-每个州distribution 差异很大 （尤其是2nd derivative 会fluctuate） (Thomas

* Mobility (data) more exploration (Jason)

4.Analysis of current models & techniques (PX

* LSTM hard to stabilize, training result vary a lot (with state, between)
* LSTM excellent result for some states
* Using smoothing helps model convergence
* Currently, using just the predicted as predictor performs the best (single variate)
* Best achieved by linear, but could tell that overfit and bottlenecked around 2, but poly can sometimes achieve better result

5.Timeline of future plans (Thomas

* Try out polynomial improve
* Try GLM
* Incorporate mobility data (after some processing)
* Specify, train for window size (hyper parameter)
* Mention current result (Kaggle score)