Warfarin Dosage with LinUCB and Supervised Learning

Haojun Li

Department of Computer Science, Stanford University

Abstract

Predicting the correct dosage of Warfarin is a challenge. Most researchers estimate the correct dosage by first gathering data on the correct dosage and then eventually devise an algorithm for the correct dosage. [1] We experimented with the following methods

- LinUCB with Disjoint models. Predicting the reward for each dosage category independently
- Supervised learning with linear models in a RL-fashion

We showed that both models are able to beat the Fixed Dose baseline and Clinical Dosage baseline in terms of regret, but unable to beat the Clinical Dosage baseline in terms of fraction of incorrect cases.

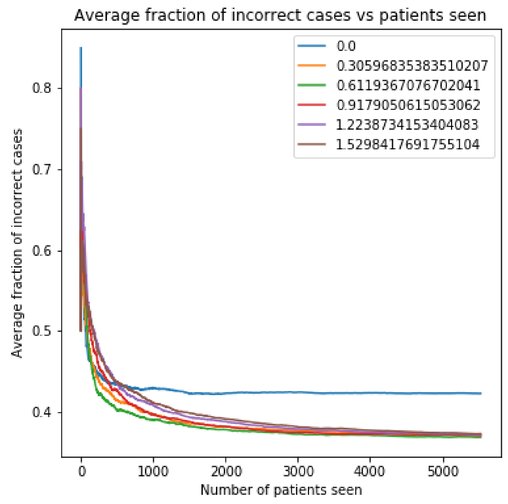
Data and Features

The data we were given contains patient features (such as age and height) as well as the actual Therapeutic Dose of Warfarin. The goal is to predict the correct dosage of Warfarin from these features using a linear model in a simulated setting where patients comes in one at a time.

One major assumption is that the actual dosage of Warfarin is linear to the features of the patients. In order to keep comparisons against the baselines fair, we will use the same set of features that Clinical Dosage baseline uses, which are **Race, Age in decades, Height (cm), Weight (kg), Carbamazepine (Tegretol), Amiodarone (Cordarone), Pheny- toin (Dilantin), Rifampin or Rifampicin.** These are also the features used in computing the regret for each algorithm

LinUCB α search

The LinUCB algorithm with disjoint models is the same as referenced in the news recommendation paper [2]. However, the parameter α needs to be tuned to push the agent to explore less and exploit more due to the short time horizon (5528 patients)



Supervised Learning Approach

In a supervised setting, the agent will see 50 patients at a time and train a model on all the data gathered as well as their true dosage. Then at each step, the agent will predict which dosage category should be used. Table below shows 2 different model classes when fitted and evaluated on the entire dataset (thus an upper limit of performance)

Model Class	Optimal Regret	Optimal Fraction
Logistic	**30.93**	**0.3538**
SVM	61.53	0.3882

Table 2:Results from different model classes

SVM does significantly worse than Logistic Regression since there is linear structure in the dataset.

Results and Discussion

We evaluated all the models using 20 different trials. For LinUCB, we used $\alpha = 0.612$ and for Supervised Learning approach we used a batch size of 50. Results are in table 1

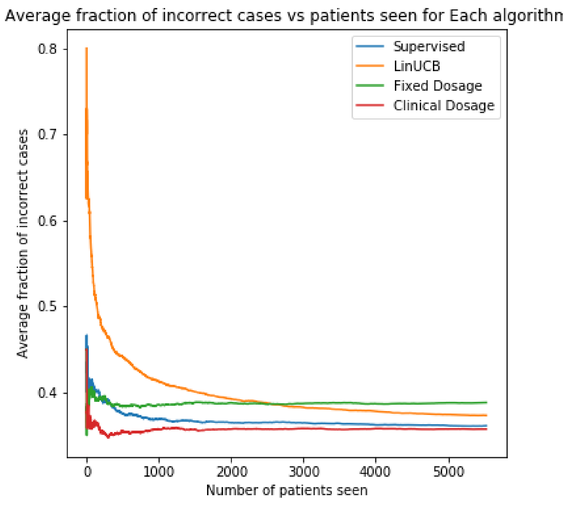
Method	Avg Total Regret	Avg Fraction
Fixed Dose	61.53	0.3882
Clinical Dosing	54.227	**0.3572**
LinUCB	51.45	0.3692
Logistic	**33.49**	0.3598

Table 1:Results from experiments

LinUCB and Supervised Learning does better in terms of Fixed Dose and Clinical Dosing Algorithms in terms of regret but not able to achieve better average fraction of incorrect cases due to the fact that it has an initial exploration period.

The plot below shows that LinUCB does explore longer than supervised learning before finally arriving at a good set of parameters and quickly catching up to the other models, while Clinical Dosing Algorithm makes correct predictions right from the beginning

Plots



Conclusion

- LinUCB with a lower α value is able to achieve lower regret than both baselines. This low α balances exploration and exploitation, and attempts to arrive at a sub-optimal parameters quicker due to the short horizon
- Supervised Learning method performed significantly better than other methods, but due to its exploration period it is unable to outperform Clinical Dosing Algorithm in terms of average fraction of incorrect cases.

Future Works

A lot still need to be done if I have more time

- Explore different Bandit algorithms with better bounds.
- Explore more α values for a better balance in exploration/exploitation

References

[1] International Warfarin Pharmacogenetics Consortium. Estimation of the warfarin dose with clinical and pharmacogenetic data. *New England Journal of Medicine*, 360(8):753–764, 2009.

[2] Lihong Li, Wei Chu, John Langford, and Robert E Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 661–670, 2010.

[3] Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

[4] Tor Lattimore and Csaba Szepesvári. Bandit algorithms. *preprint*, page 28, 2018.

[5] Varsha Dani, Thomas P Hayes, and Sham M Kakade. Stochastic linear optimization under bandit feedback. 2008.