Estimation of the Warfarin Dose

CS234 Winter 2023 Final Project

Tony Cai (tonycai@stanford.edu)
Jiatu Liu (ljiatu@stanford.edu)

Introduction

Background

Warfarin is the most widely used oral blood anticoagulant agent worldwide, but it's difficult to establish the correct dosage because it can vary substantially among patients, and the consequences of taking an incorrect dose can be severe.

Our Work

Empirically evaluate the effectiveness of various linear contextual multi-arm bandit algorithms in terms of overall accuracy and safety.

Dataset

- PharmGKB Warfarin dosage dataset
  - Patient features: biological (race, weight, height), history of medicine intake, genetic
  - Physician dosage
- Preprocessing
  - Remove patients who didn't reach stable dosage
  - Impute missing features

Methods

- Baselines
  - Fixed dosage: assign 35 mg/wk (medium) dose to all patients.
  - Linear regression (Phamacogenetic dosing algorithm)
  - "Bandit" version of supervised learning algorithm

Algorithm 2 "Bandit" Version of Supervised Learning

Initialize a linear predictor $f_0(x)$ randomly
for $t = 1 \dots N$ **do**
  Observe patient feature x_i
  $y_i \leftarrow f_{i-1}(x_i)$ ▷ Predict the dosage using the current predictor
  Refit new predictor $f_i(x)$ using $x_{1:t}$ and $y_{1:t}$
end for

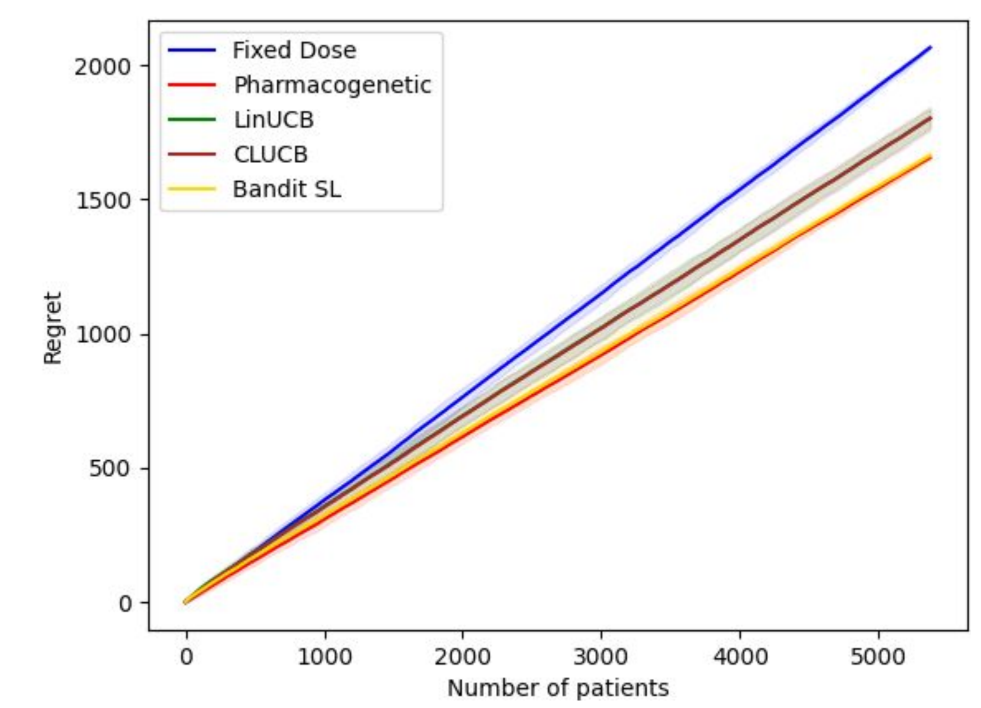
- MAB problem formulation:
  - Bandit arms: low dosage (< 21 mg/wk), medium dosage (>= 21 and <= 49 mg/wk) and high dosage (> 49 mg/wk)
  - Bandit only observes whether its chosen action is correct
  - $\theta^T = [\theta_1^T \quad \theta_2^T \quad \theta_3^T] \qquad x_{t,0}^T = [x_t^T \quad 0^d \quad 0^d] \qquad \mathbb{E}[r_t(a)] = \theta^T x_{t,a}$
- LinUCB and Conservative LinUCB

Algorithm 1 Conservative LinUCB (CLUCB)

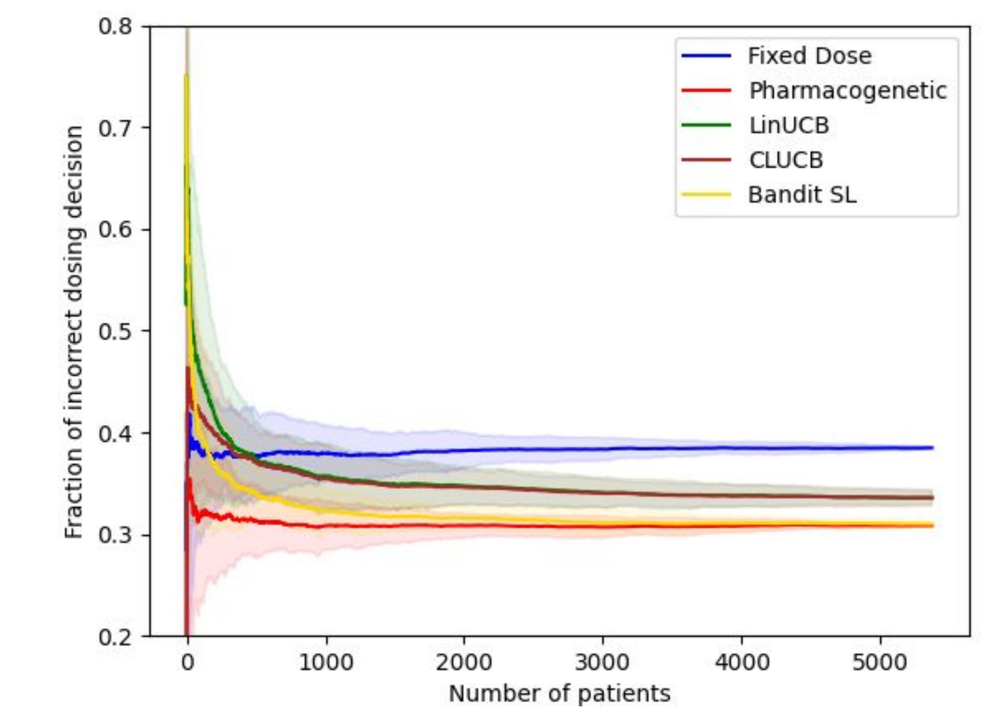
Require: Confidence bound β_t
Require: Maximum acceptable performance degradation α
$\hat{\theta} \leftarrow 0^d$
$V \leftarrow \lambda I$
for $t = 1 \dots N$ **do**
  for $a \in \mathcal{A}$ **do**
    $p_{t,a} \leftarrow \hat{\theta}^T x_{t,a} + \beta_t \sqrt{x_{t,a}^T V^{-1} x_{t,a}}$ ▷ $\mathrm{UCB}(a) \leftarrow \max_{\hat{\theta} \in C_t} x_{t,a}^T \hat{\theta}$
  end for
  $a' \leftarrow \arg\max_a p_{t,a}$
  $z \leftarrow \sum_{t \in S} X_t$
  $L \leftarrow z^T \hat{\theta} z - \beta_t \sqrt{z^T V^{-1} z}$ ▷ $L \leftarrow \min_{\hat{\theta} \in C_t} \sum_{t \in S} x_{t,a}^T \hat{\theta}$
  if $L + \sum_{t \in \bar{S}} r_b \geq (1 - \alpha) R_b$ **then**
    $\hat{\theta} \leftarrow (X_{1:t}^T X_{1:t} + \lambda I)^{-1} X_{1:t}^T Y_{1:t}$ ▷ Compute $\hat{\theta}$ using linear regression
    $V \leftarrow \lambda I + \sum X_t X_t^T$ ▷ Update $C_t = \{\theta \in \mathbb{R}^D : ||\theta - \hat{\theta}||_{V_t} \leq \beta_t\}$
    Take action a'
  else
    Take action a_b
  end if
end for

Results

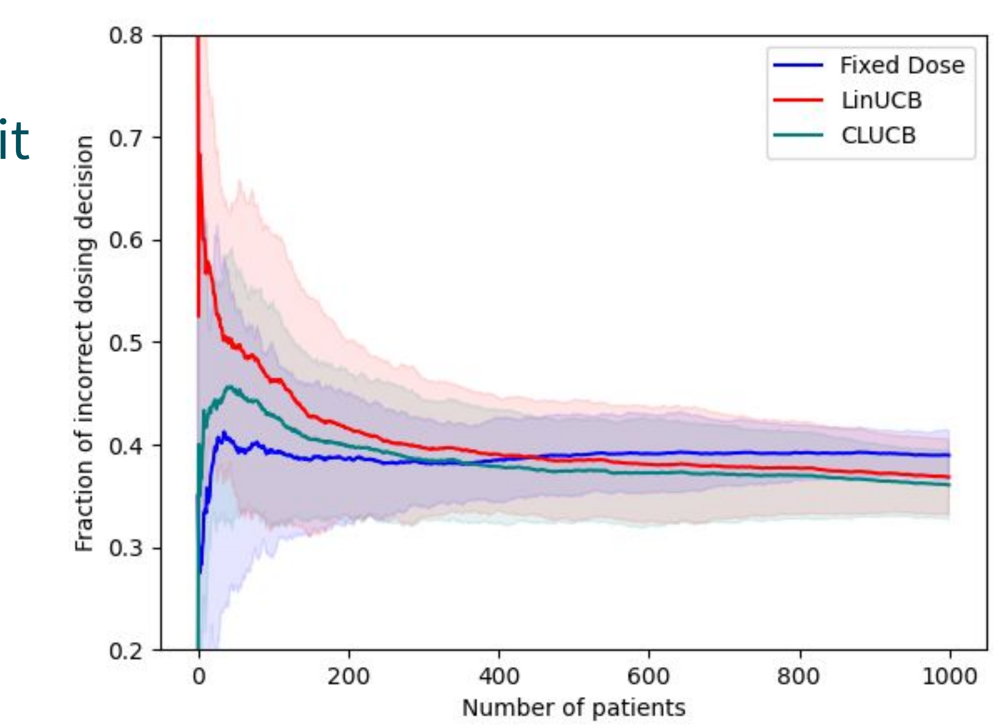
Regret



Incorrect Dosage Fraction



- LinUCB performs better than baseline but worse than SL bandit
- Bandit performance limited by linear model and features
- CLUCB vs LinUCB
  - Makes significantly fewer mistakes on initial patients
  - No adverse effect on overall performance



Future Work

- Explore more sophisticated confidence interval construction
- Explore non-linear models and Thompson Sampling approaches
- Minimizing number of catastrophic decisions (e.g. by adjusting reward structure)

References

- Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation.
- Abbasi-Yadkori, Yasin, Dávid Pál, and Csaba Szepesvári. "Improved algorithms for linear stochastic bandits."
- Abbas Kazerouni, Mohammad Ghavamzadeh, Yasin Abbasi Yadkori, and Benjamin Van Roy. Conservative contextual linear bandits.