

# Framework to Recommend Treatments for Lyme Disease

Eric (Kaiyuan) Chen<sup>1</sup>, Rong Huang<sup>1</sup>, Diyi Liu<sup>2</sup>, Catherine Wahlenmayer<sup>3</sup>, Ada (Jiewen) Wang<sup>1</sup>  
<sup>1</sup>University of California, Los Angeles, <sup>2</sup>Shanghai Jiao Tong University, <sup>3</sup>Gannon University

## Overview

Lyme disease is greatly understudied; however, patients have banded together to pool their data via the MyLymeData survey through LymeDisease.org to facilitate research. We created an adaptable framework to recommend antibiotic and alternative treatments based on this patient data. The framework can also predict potential side effects from alternative treatments. It utilizes relevant portions of patient survey data to find medical regimens for unwell patients based on patients with similar experiences. The framework is adaptable to different questions within the context of MyLymeData or similar contexts as well as different classification methods.



## Data

The MyLymeData is built from patient survey responses about their illness. The embedded data are broken down based on whether the patient has recovered (400 Well) or not (3600 Unwell) from Lyme disease.

The Well and Unwell classes are very unbalanced. To counter this, we

- Under-sample by truncating the Unwell
- Under-sample by random selection of Unwell
- Over-sample by duplicating the Well
- Over-sample by appending the average of two random Well rows

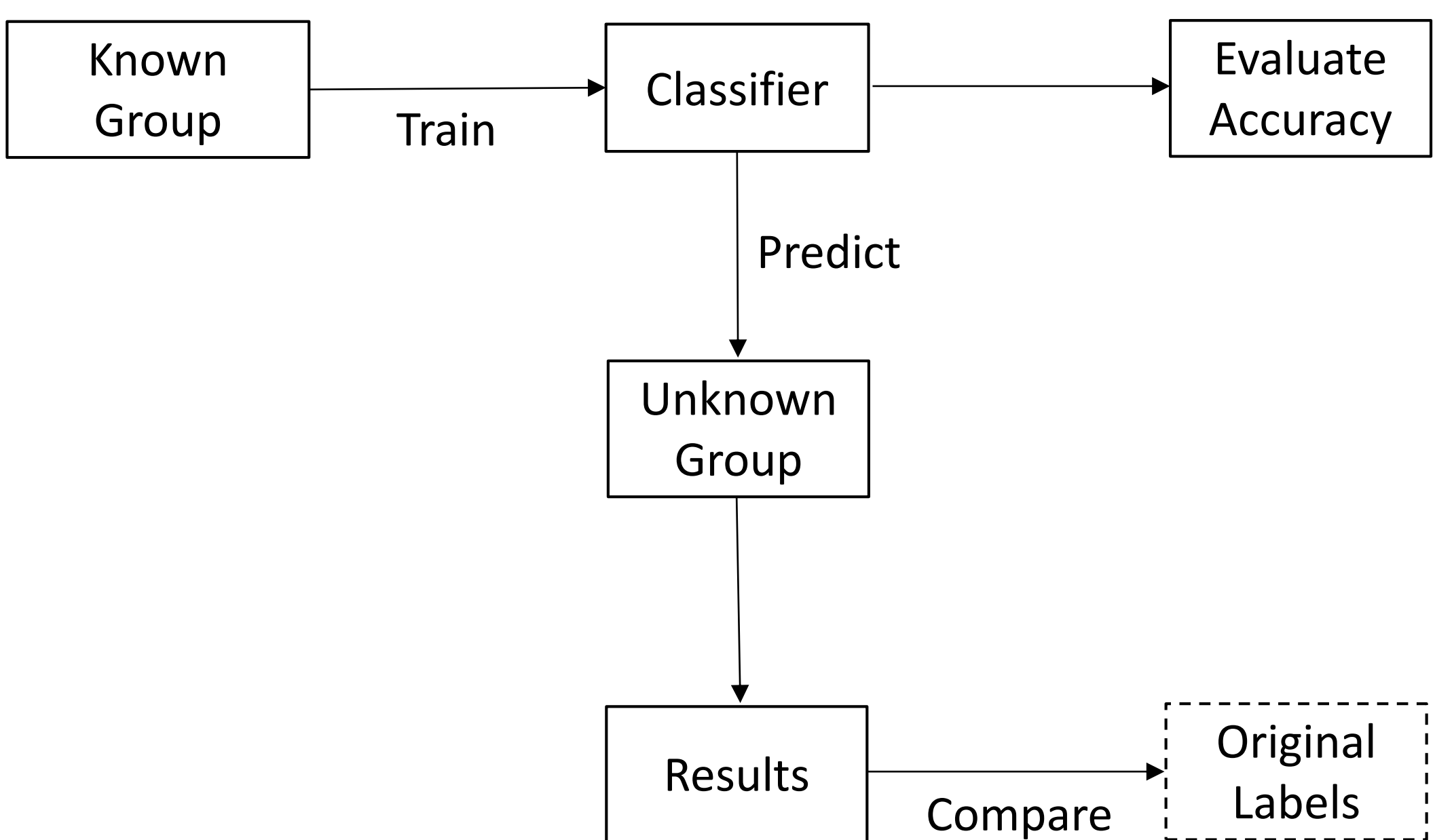
## Questions of Interest

The following questions motivated our work:

1. **How effective will an antibiotic be?**
2. What is the most effective alternative treatment?
3. What level of side effects will a patient experience form an alternative treatment?

We focused on question 1 and designed a recommender framework that predicts the effectiveness of antibiotics and then compared these predictions with the actual antibiotic regimen of each patient.

## Framework



We implemented the recommender framework above using neural networks and simple binary classification. We use it to recommend antibiotics, predict the level of side effects patients will experience, suggest exercise protocols, etc.

We used the following classifiers on the MyLymeData.

- **Neural networks:** multilayer perceptron network; in the framework, we use a fully connected network using the Python package Pytorch
- **Simple Binary Classification** [1]: classifies each binary data point using a sign pattern that describes on which sides of a selected number of random hyperplanes the point lies, then assigns the point to the class that has the matching sign pattern most often

To predict the **effectiveness of various antibiotics** for Unwell patients from Well patients who share similar conditions:

- Train using the **Well** baseline questions
- Label the antibiotics being taken
- Predict based on **Unwell** baseline questions
- Label potential effectiveness of future antibiotics

We present an example of the application of this framework in the table on the right.

## Results for Question 1

We applied the framework to give insight to Question 1 first, “How effective will an antibiotic be?”

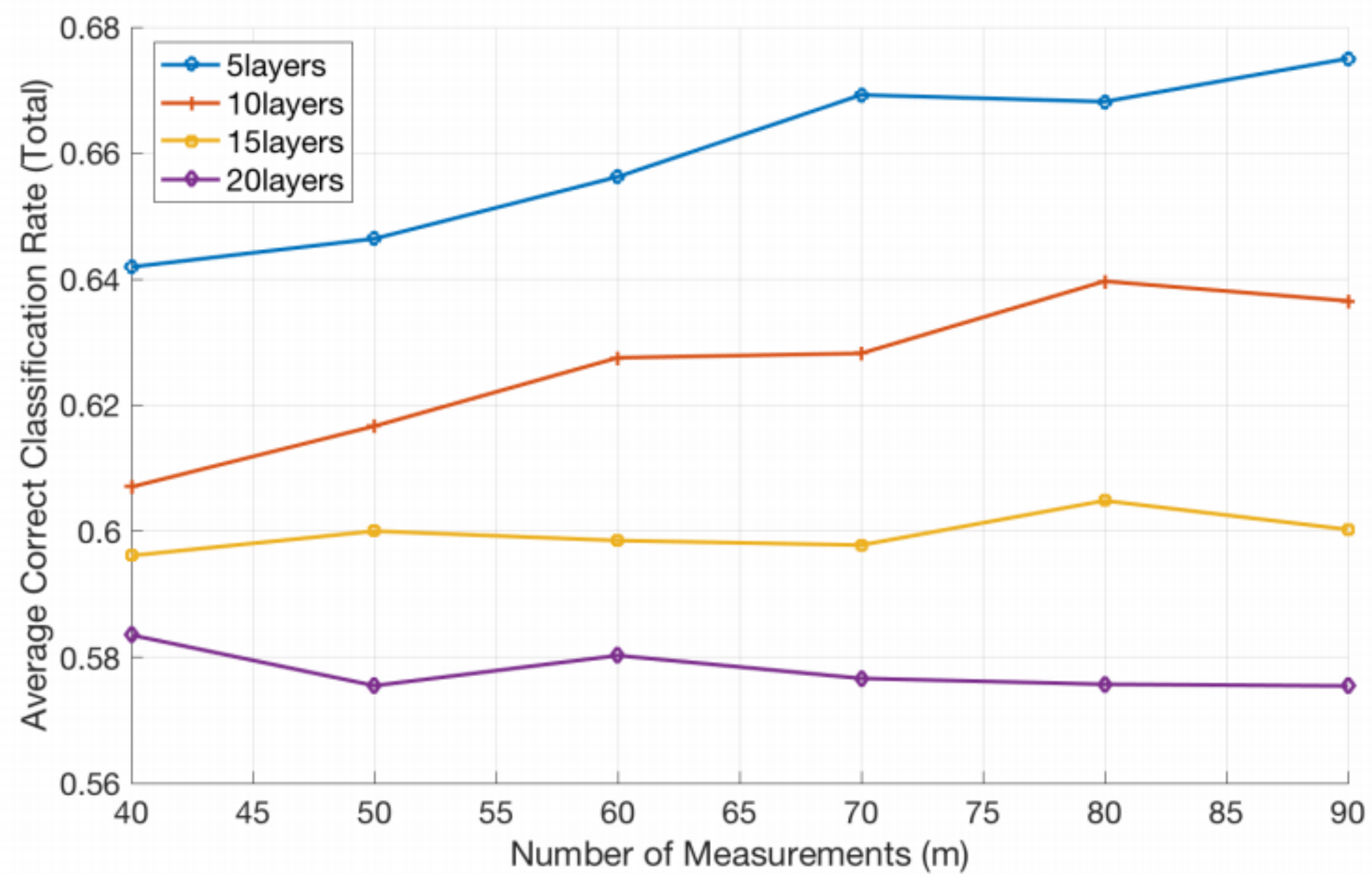
Patients’ current antibiotics compared to predicted **match completely for 34%** of patients. **22% differ by one** out of 37 antibiotics, **14% differ by two**, and **only 30% differ by more than two**. The suggests that our model is reliable and matches real-world data while not overfitting.

## Results

A sample of synthetic data for question 1 is in the table. Matching rows indicate the reliability of the model. Rows that do not match, especially for instances in which the patient did not answer, allow us to predict the effectiveness of the antibiotic. In this example, we would recommend the individual to consult a doctor about the antibiotic Biaxin because the model found that it was effective for patients with similar baseline question responses.

Antibiotic	Original patient data	Model prediction	Interpretation
Alinia	1	1	Agreed (effective)
Amoxicillin	1	1	Agreed
Amoxicillin clavulanate	-1	0	Unanswered, predict ineffective
Biaxin	-1	1	Unanswered, predict <b>beneficial</b>
Cedax	-1	0	Unanswered, predict ineffective
Cefuroxime	0	0	Agreed (ineffective)
...			(randomly generated sample data)

This figure shows the accuracy of the framework using simple binary classification to evaluate the effectiveness of antibiotics. It assumes Well and Unwell patients’ antibiotics are effective and ineffective respectively. The data is balanced by random under-sample and accuracy is evaluated using 10-fold cross-validation. Ordinarily, increased layers increases performance, but 5 layers is best here.



## Future Work

Evaluate the accuracy of the recommender using other classifiers such as support vector machines, random forests, and logistic regression. We could also combine some of these using ensemble learning to try to improve performance.

### References

[1] D. Needell, R. Saab, and T. Woolf, “Simple classification using binary data,” CoRR, vol.abs./1707.01945, 2017