

The slide features abstract geometric designs. The top-left corner is composed of several overlapping triangles in shades of blue, green, and red. The bottom-right corner features a cluster of overlapping triangles in various shades of gray.

Drug Review Analysis

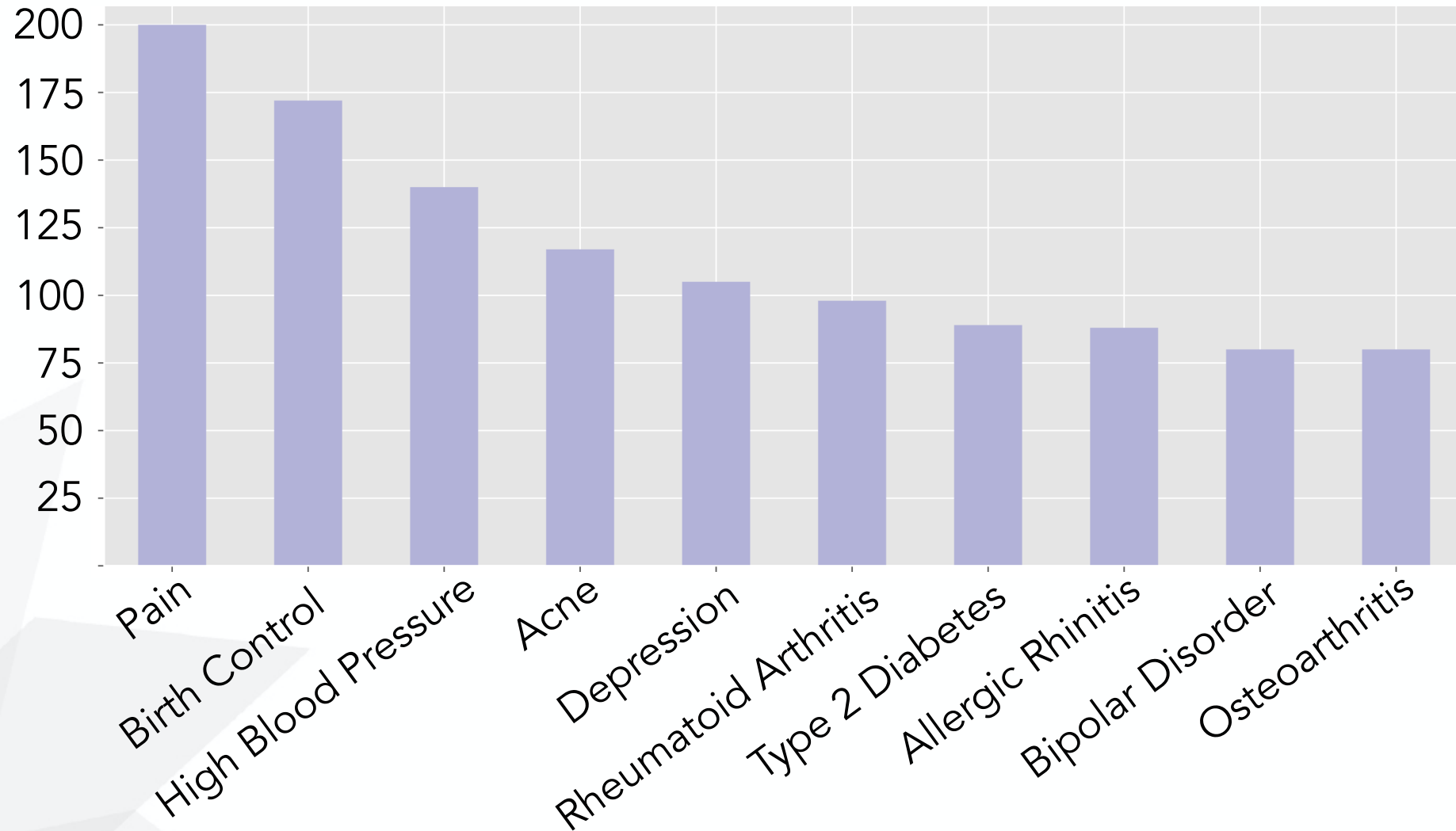
A vertical decorative element consisting of three thin, parallel stripes in orange, yellow, and blue.

JiEun Song



What makes drug
reviews useful?

Top 10 Conditions with Most Number of Drugs



Problem in the Market

- Too many different drugs for a condition
- Patients often have very little information

➔ Help patients' decision based on their concerns

Workflow

Data
Collection



UCI machine
learning
repository

Preprocessing



Clean data
Remove stop words
Lemmatizing
CountVectorizer

Topic
Modeling



Latent Dirichlet
Allocation

Visualization



Python
Tableau

Topic Modeling

Side Effect

year old dry mouth
lost lb no side
side effect
lost pound blood pressure
started taking
year ago
weight loss

Feeling

took pill
feel like within hour
felt like
hour later
unprotected sex
first time
yeast infection
two day first dose

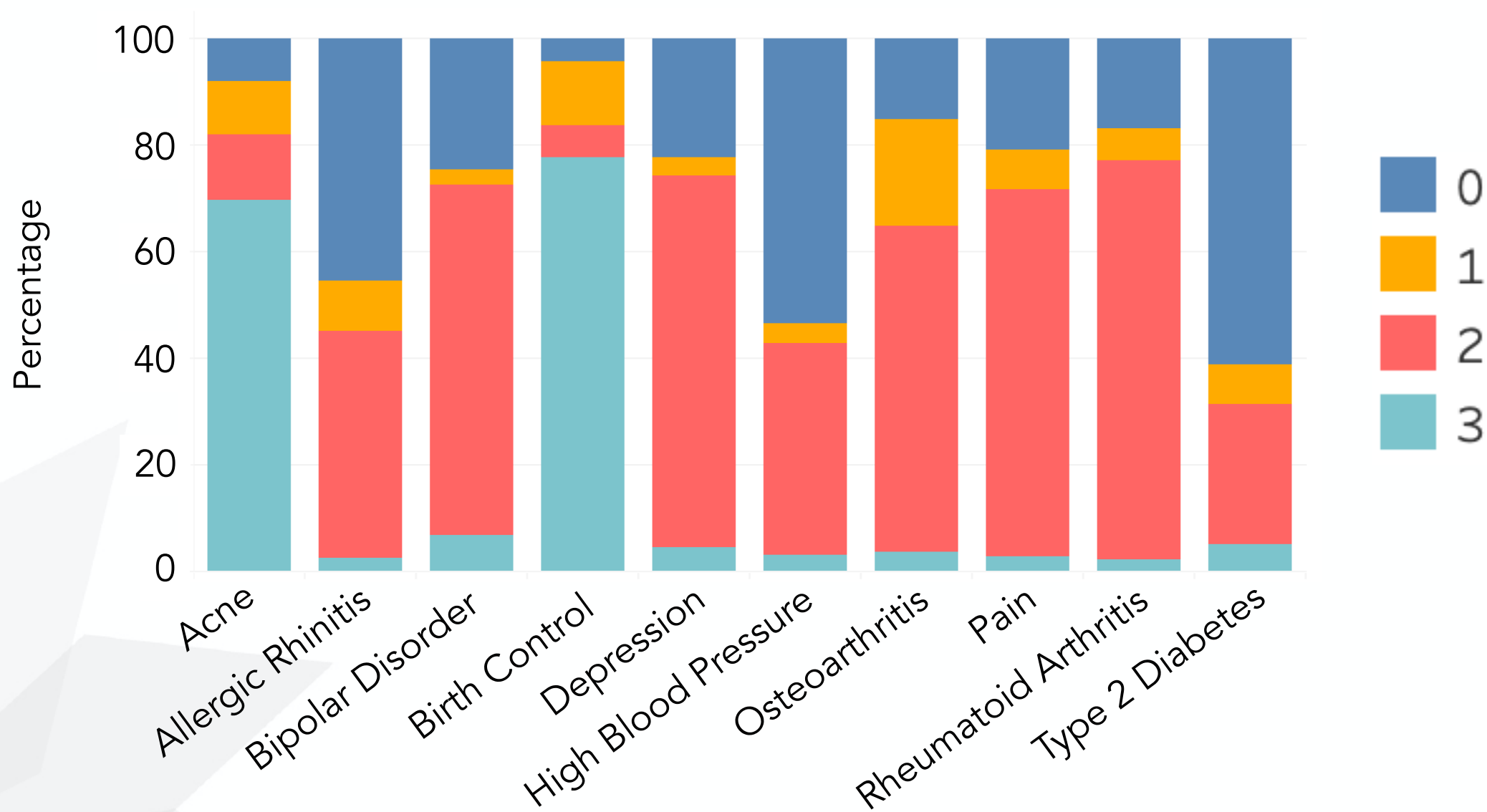
Efficacy

started taking
panic attack
no longer
much better
depression anxiety could not
year old year ago
back pain
feel like

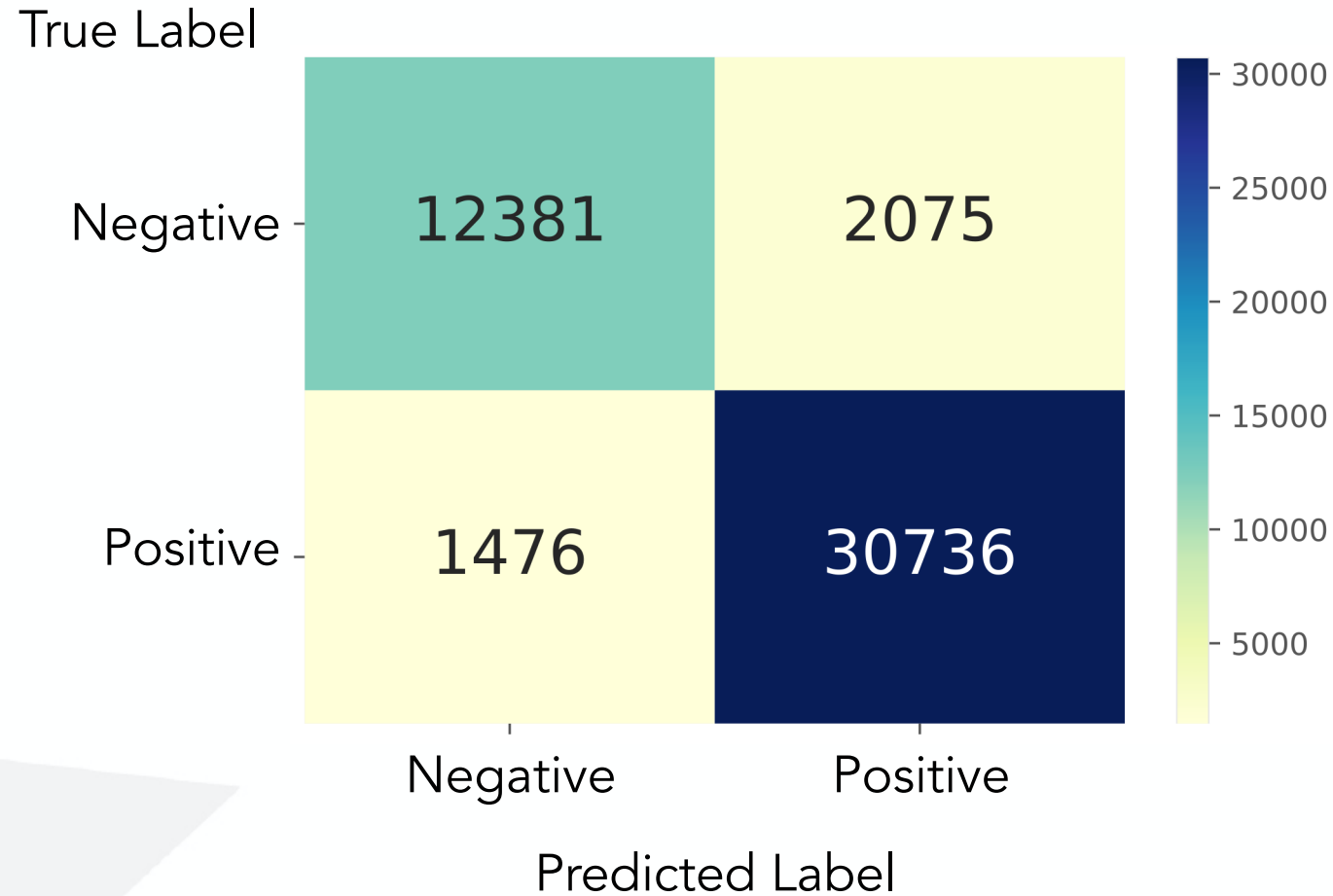
Birth Control

gained weight
no weight **weight gain**
taking pill
birth control
first month
mood swing
sex drive
side effect
gained pound

Topic Distribution of Top 10 Conditions



Predicting Rating – Logistic Regression



Accuracy Score 0.92

Business Value and Future Work

- Build a recommender system:
 - Help patients' decision making
 - Provide benchmark to drug providers
- Topic modeling for reviews across different conditions



Thank you!

Question?