

PREDICTING DRUG-DRUG INTERACTIONS: COMBINING MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING

Alexander Hayes — Savannah Smith — Ciabhan Connelly — Devendra Dhami — Sriraam Natarajan
hayesall@indiana.edu — savannah.smith@valpo.edu — morcachomrin@gmail.com — ddhami@umail.iu.edu — natarasr@indiana.edu

SCHOOL OF INFORMATICS
AND COMPUTING

Indiana University | Bloomington

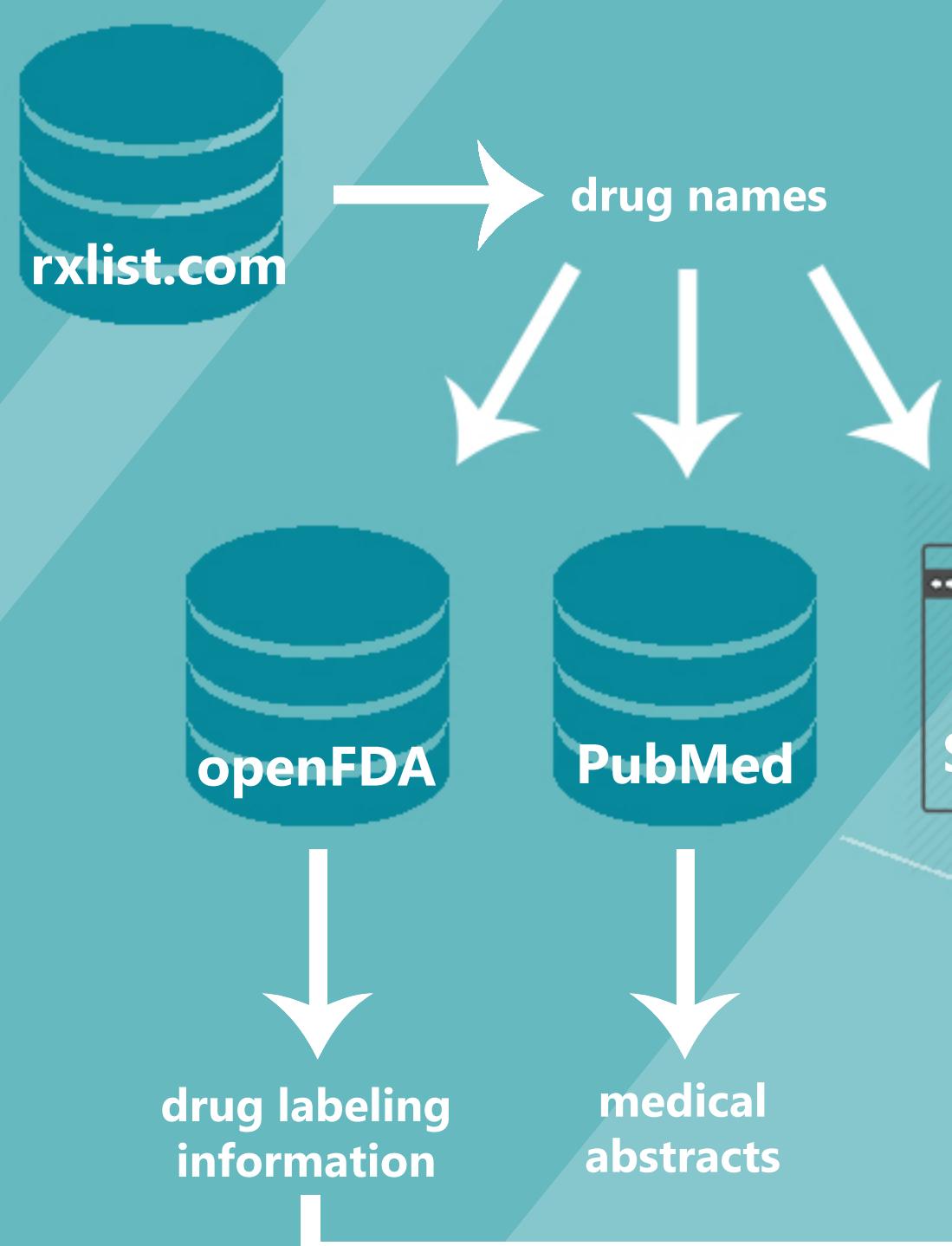
STARAI



ProHealth
Indiana University

The problem with polypharmacy

Polypharmacy: four medications taken at the same time. **44% of men and 57% of women aged 65 and older** find themselves in this situation. After four, the chance of experiencing an adverse reaction **increases exponentially**, costing **thousands of lives and millions of dollars annually**.



Creating a dataset

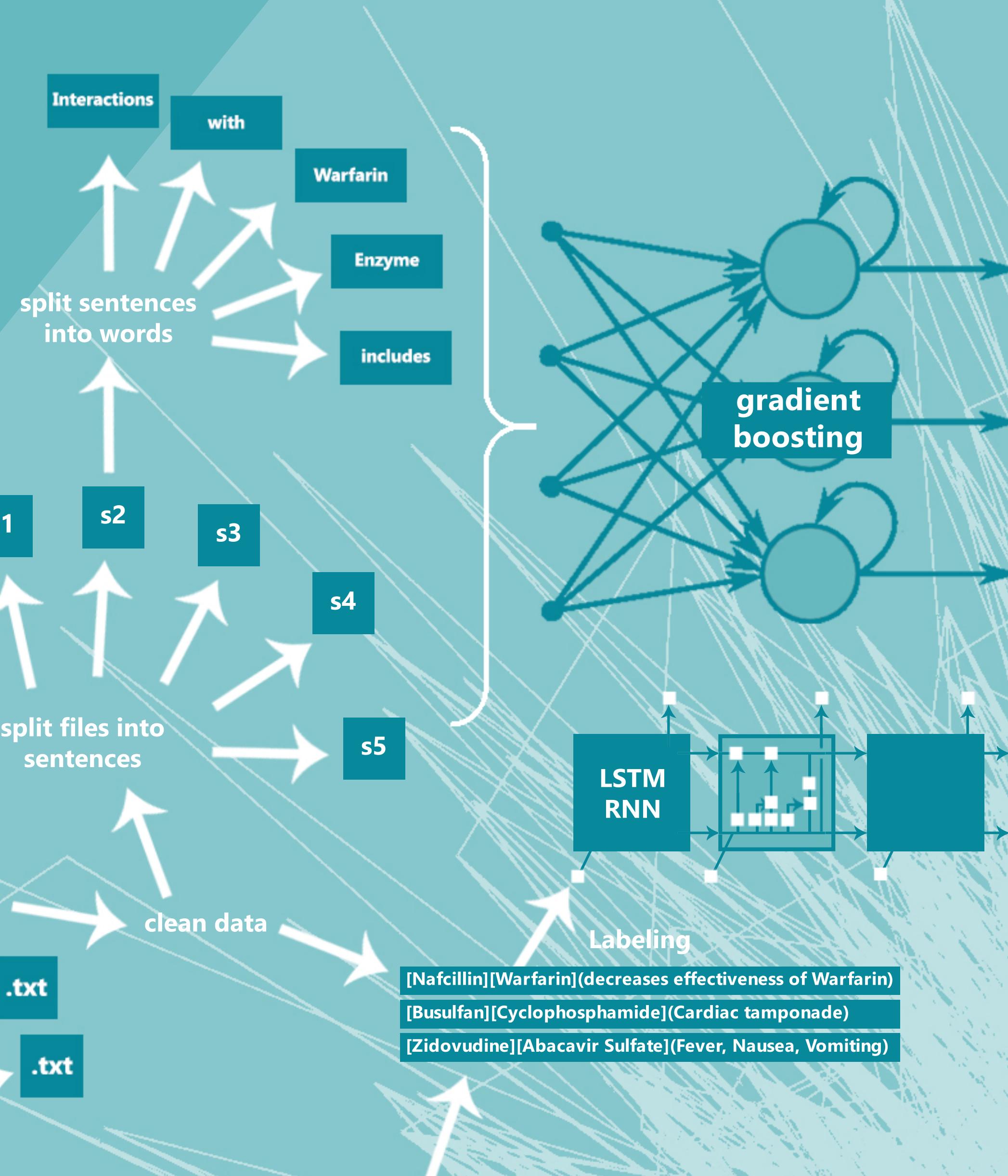
adverse reactions,
drug interactions,
warnings and cautions.

studies on drug combinations

Patient experiences with their medications

Labeling for NLP

Take **phrases from the data**: "interacts with," "inhibits cytochrome p450," "no evidence of." These become **positive** and **negative** examples that we can point to in the data. We had **250 positive examples** and **100 negative examples** for training and testing.



Building the model / Results

openFDA labels became the foundation for an **LSTM network** that estimated the **likelihood of an interaction occurring** when two drugs were taken at the same time.

LSTM produced a **73.1% accuracy**.

Training - 50% Validation - 10% Testing - 40%

Pictured in the background is the network of drug interactions generated from PubMed searches.

Further Work

More models: further training with DeepDive

More Data: train with more data sets (PubMed, family medical history, blogs)

Privileged Information: Create a "truthy" model: training on blogs with privileged info