

Medical Text Classification

StarHacks Datathon

Team Member: Caijun Qin, Jeongin Lee,
Kushal Dev, Joy Chen



What is the Problem?

Medical jargons and limited health literacy often lead to communication barriers in healthcare settings. The 2003 National Assessment of Adult Literacy found that nearly **half** of the US population has inadequate literacy.

This leads to a host of downstream effects



Inappropriate Health Decisions

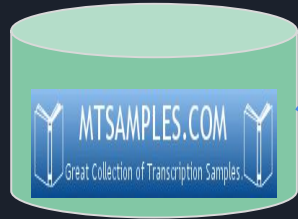
More Serious Health Outcomes

Low health literacy costs the US healthcare system up to **\$73** billion annually

Workflow

Data Preprocessing
& Model Training

Building the App &
Deployment



Data Retrieval
(from Kaggle)

IBM Z

Scikit-learn

SVM, KNN, Random
Forest



Streamlit

Prediction