



Bluenotes **GLOBAL** 2021  
VIRTUAL EXPERIENCE

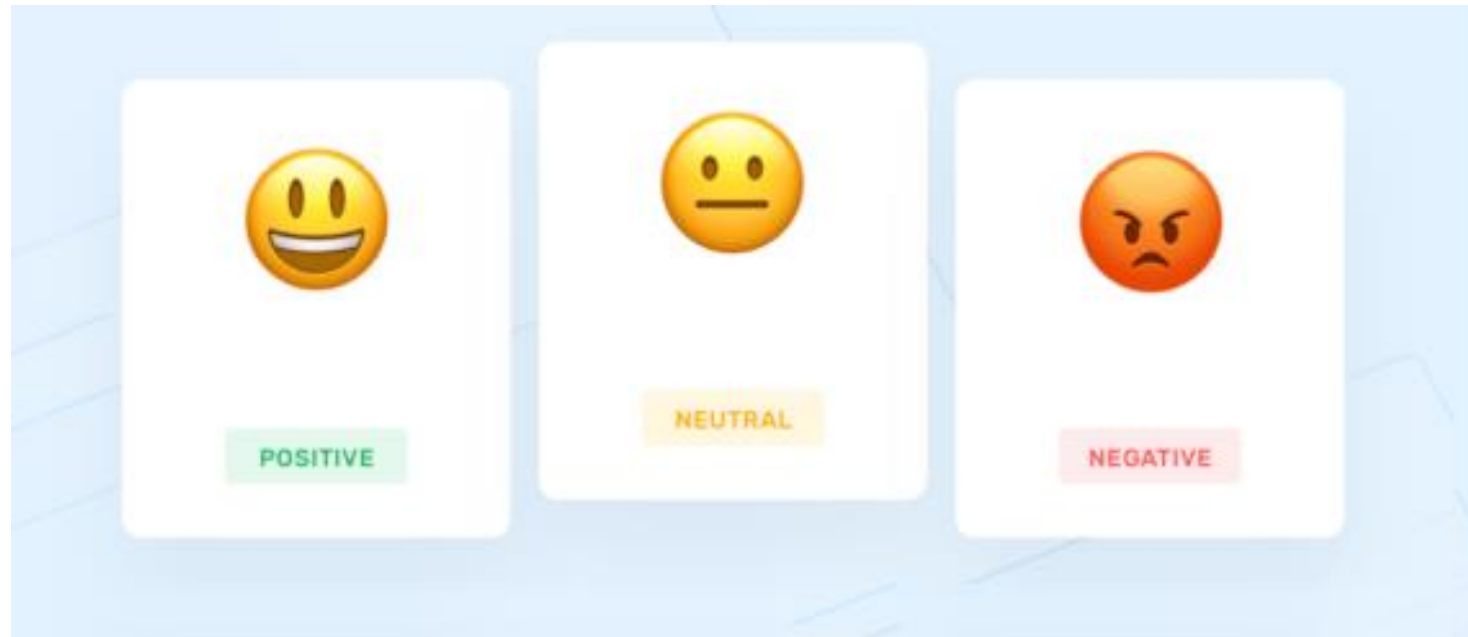
# Sentiment Analysis

Presenter: Nicole and Rahat



# Introduction

Sentiment analysis is a natural language processing technique used to determine whether data is positive, negative or neutral.



# Previous works



Product analysis



Reputation Management



Social Media Monitoring



Brand Monitoring



Market Research

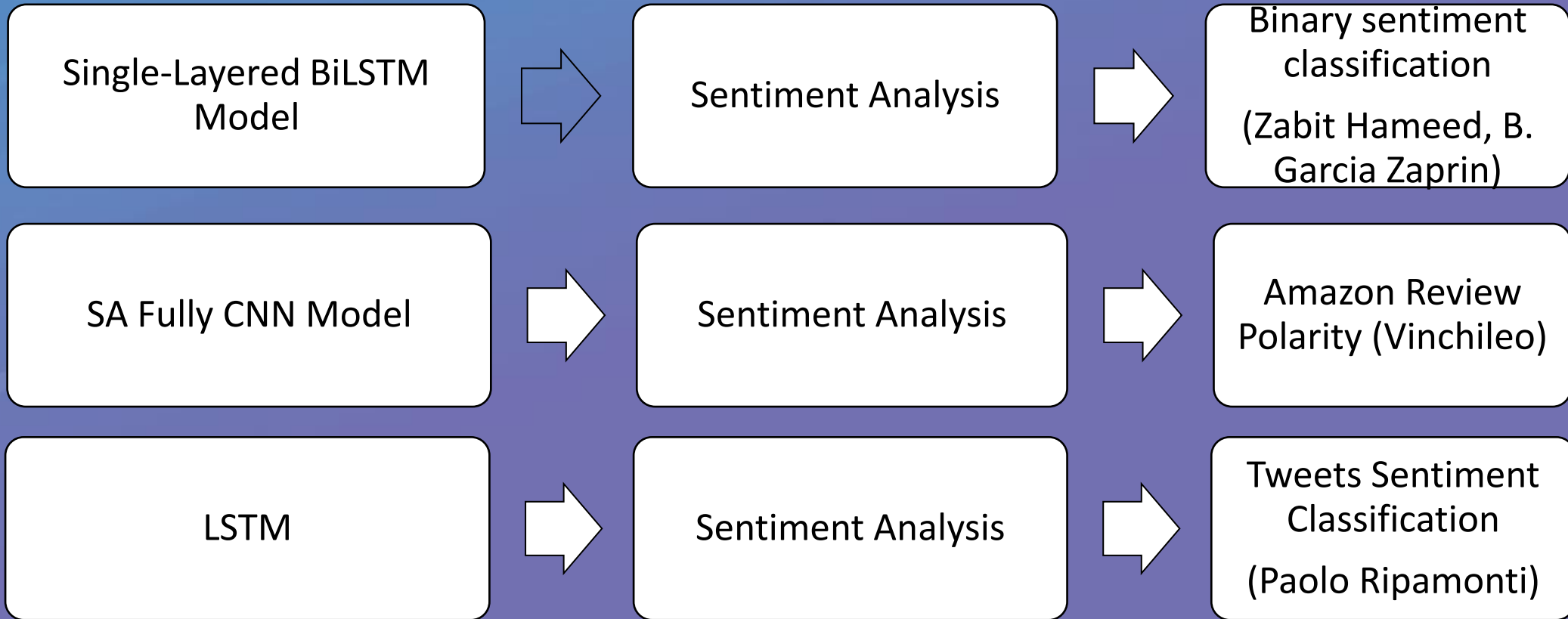


Competitor Analysis

# Student Evaluations ??



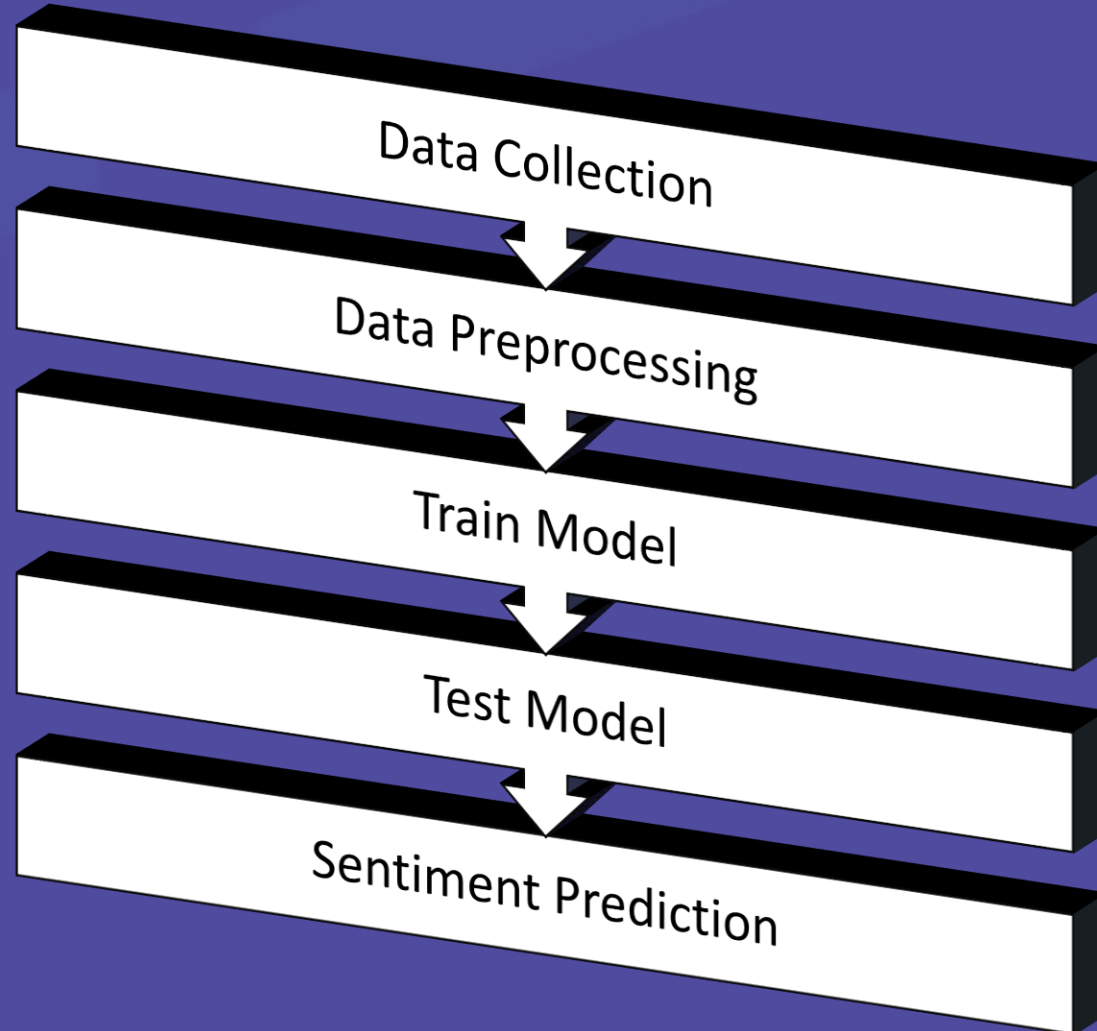
# Commonly used models



Note: LSTM stands for Long short-term memory; CNN stands for convolutional neural network



# Standard workflow





# Evaluation data collection

Label	Text
Positive	Great professor, very smart and incredibly helpful
Negative	Worst teacher I have ever had.
Positive	A good teacher. Concerned about his students.
Neutral	Nothing to say



# Evaluation data

His teaching method is WICKED!

I got a 40% on the first exam, went into office hours and the man compared accounting to golf. No matter how hard and how many times he plays golf, hell never be tiger woods...basically youre not gonnna get a hole in one!!

like that he is very tall and chinese. interesting. tallest chinese man since yao ming. only posted one of my minute papers when every one was golden material tho.

He will be missed greatly in anatomy land!!!

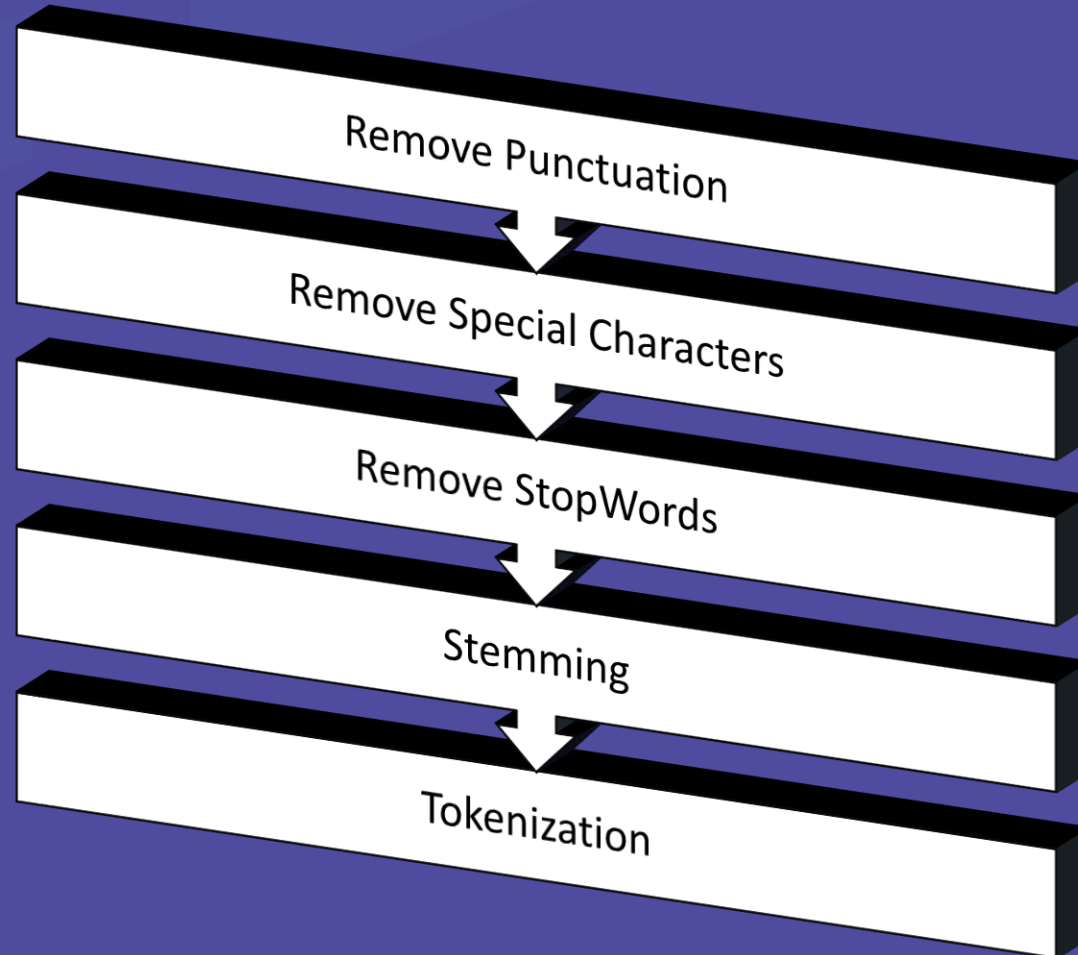
A stud!



Tricky ones!!



# Data preprocessing



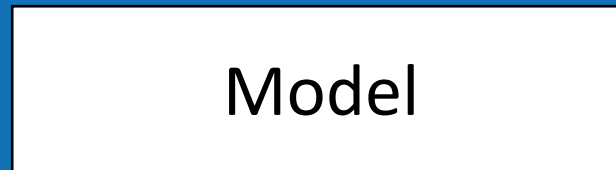




# Model training

**Train**

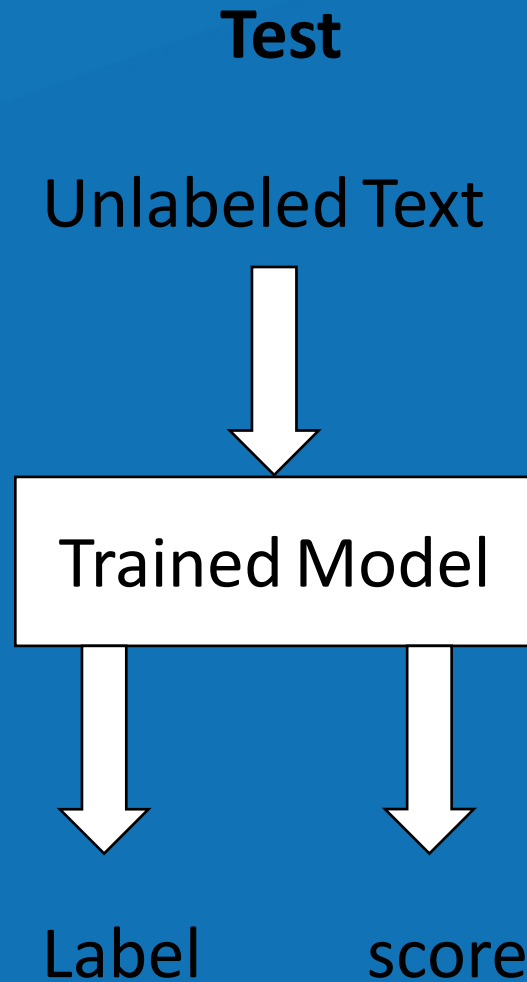
Labelled Dataset



**Trained Model**



# Model testing





# Example (model testing)

Input

Dr. Key is an amazing  
teacher with a true  
passion for Philosophy.



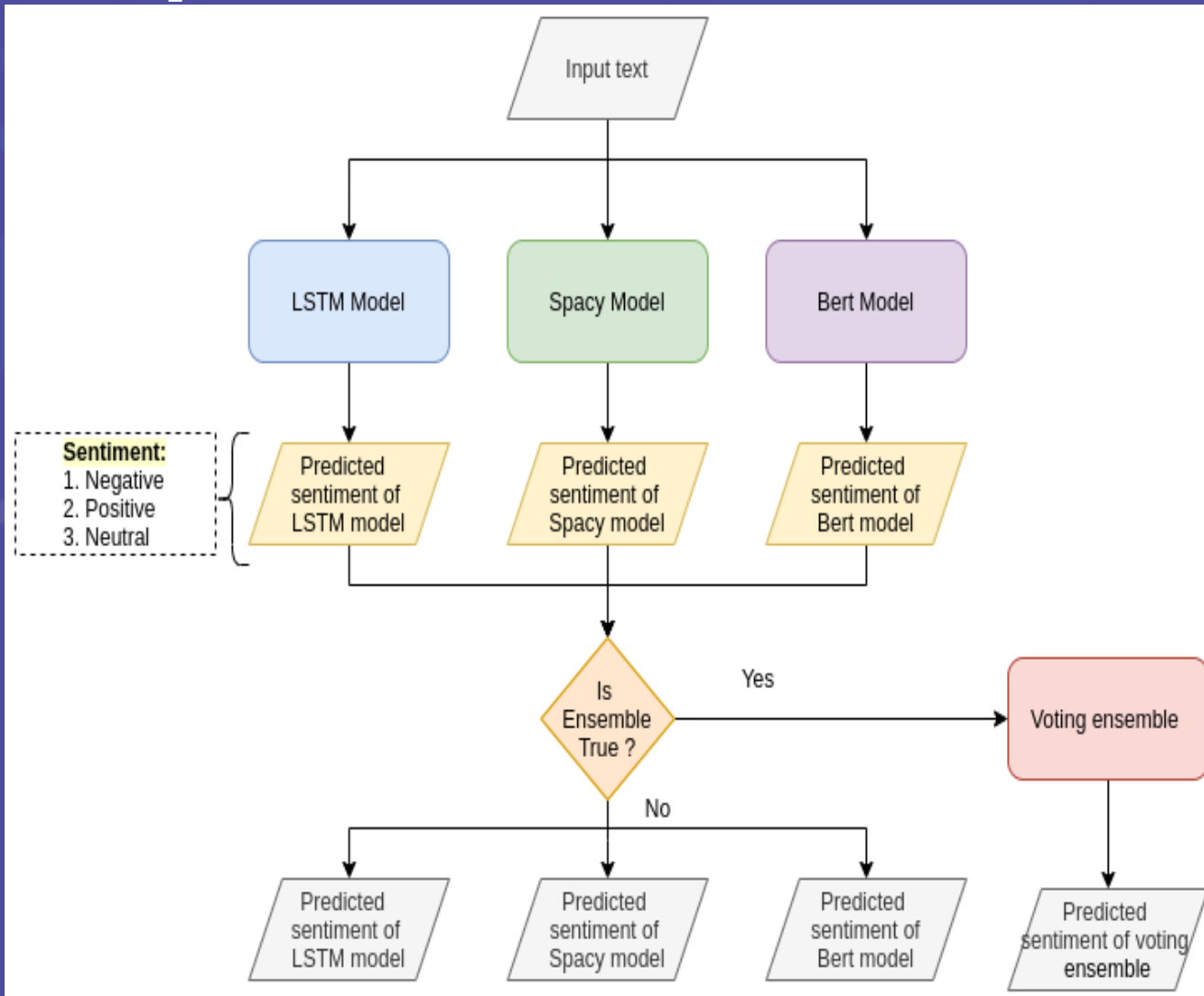
Model



Output

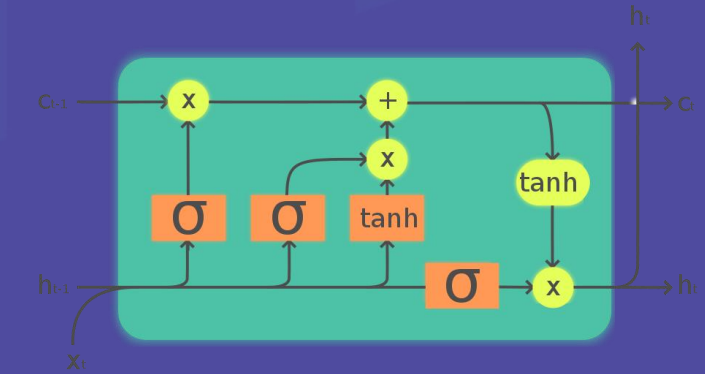
Label : Positive  
Score: 0.9789.

# Proposed framework

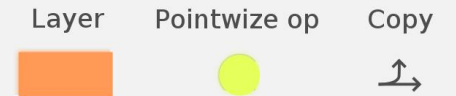


# LSTM model

- Contains a cell, an input gate, an output gate and a forget gate.
- The cell remembers values over arbitrary time interval.
- The gates regulate the flow of information into and out of the cell.

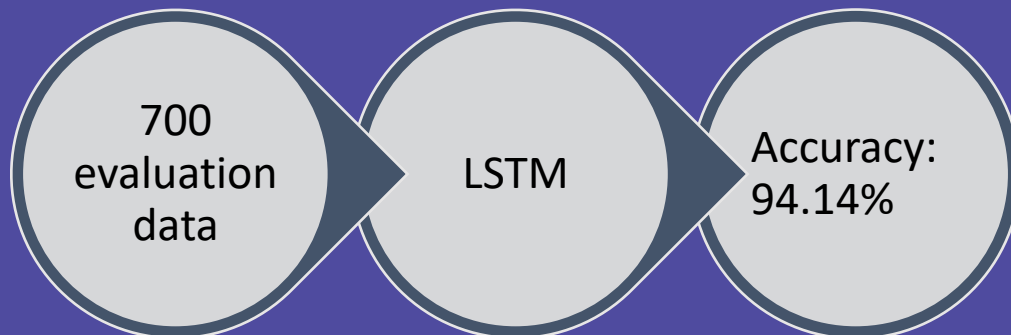


Legend:



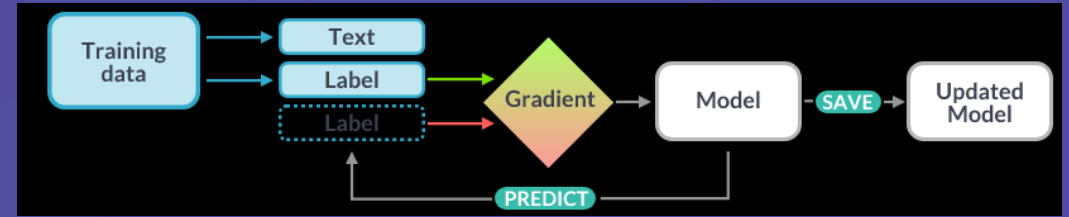
Input

Output



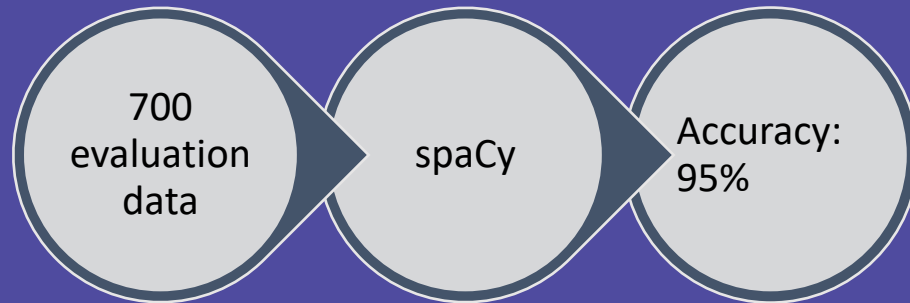


# spaCy model



- Main components are powered by statistical models.
- Every decision is a prediction based on the model's current weight values.
- Training is an iterative process.

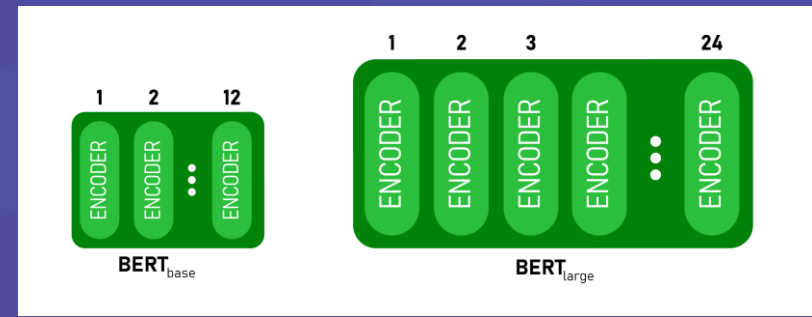
Input



Output



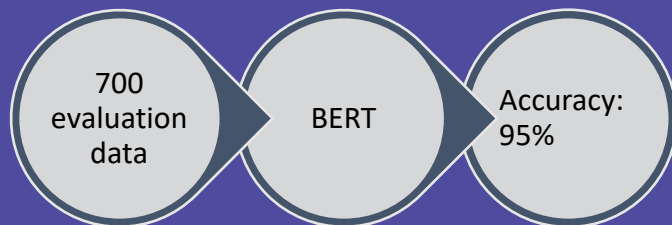
# BERT model



- Stands for Bidirectional Encoder Representations from Transformers.
- An Encoder stack of transformer architecture.
- A transformer architecture is an encoder-decoder network that uses self-attention on the encoder side and attention on the decoder side.
- BERTBASE has 12 layers in the Encoder stack while BERTLARGE has 24 layers in the Encoder stack.

Input

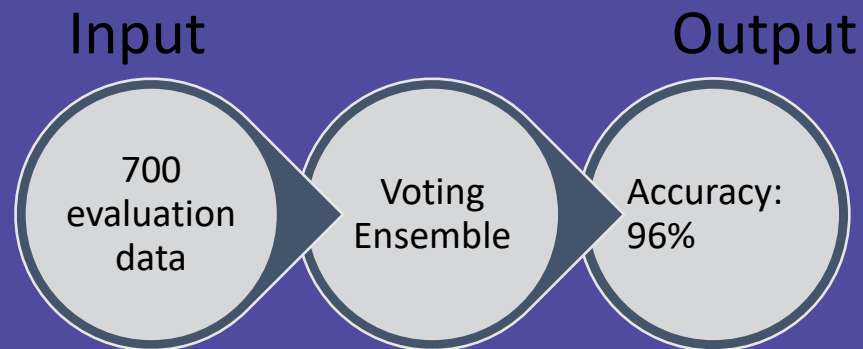
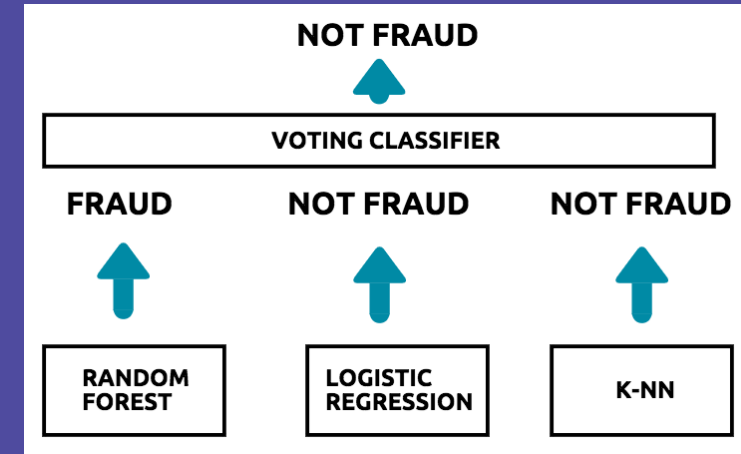
Output





# Voting Ensemble

- Ensemble machine learning model that combines the predictions from multiple other models.
- Classification method: Predictions are the majority vote of contributing models.







# Future work

- **A web application for automating the classification of survey data.**
- **Inclusion of evaluation data in the paper outline to improve students' workload management.**