6th-12th Grader

Predicting Effective Argument

Koh Way Keat

Problem Statement

- Georgia State University and the Learning Agency Lab would like to improve the effective writing skills of students.

- The education system did not put much emphasis in persuasive writing, which may hinder critical thinking development.

Objective:

- Build a model to classify argumentative elements in student writing as “effective”, “adequate” or “ineffective”.

Data

Essay ID: 331CA007D0AD

When people ask for advice, they sometimes talk to more than one person. **Adequate - Lead** Seeking multiple opinions can help someone make a better choice. Because if someone ask for advice to multiple people thats good. **Adequate - Position** Because one person might have better advice than some other person they asked if they asked multiple people for advice **Effective - Claim** . Also you can ask for advice when you need help or you don't know how to do something **Adequate - Evidence** Also asking muliple people for advice can help people make a better choice becauce other people might have diffrent advice or diffrent expereances with thething the person is asking them. **Effective - Claim** Thats one reason why asking more than one person can help people make a better choice **Ineffective - Evidence** Another reason why asking multiple people for advice is good is that some people might give you more advice then others. **Adequate - Claim** A reason someone might not ask for advice might not want to get advice beacuse they might think it will make them look weak for asking for advice. A second reason someone might not ask for advice might be beacuse their nervous to ask. A third reason might be because they might won't have that much people they can ask for advice. **Adequate - Counterclaim** When poeple ask for advice they might sometimes ask multiple people for advice. Thats good beacuse one person might have more advice then someone alse. Also someone might have better advice then the first person you asked or the second person you asked. Also asking more then one person for advice is good because some ither persons might have diffrent expereances with the thing your asking for. **Adequate - Concluding Statement**

Data - Discourse Type

Discourse Types

The dataset presented here contains argumentative essays written by U.S students in grades 6-12. These essays were annotated by expert raters for discourse elements commonly found in argumentative writing:

1. **Lead** - an introduction that begins with a statistic, a quotation, a description, or some other device to grab the reader's attention and point toward the thesis
2. **Position** - an opinion or conclusion on the main question
3. **Claim** - a claim that supports the position
4. **Counterclaim** - a claim that refutes another claim or gives an opposing reason to the position
5. **Rebuttal** - a claim that refutes a counterclaim
6. **Evidence** - ideas or examples that support claims, counterclaims, or rebuttals.
7. **Concluding Statement** - a concluding statement that restates the claims

01

Data Cleaning & EDA

Data Cleaning

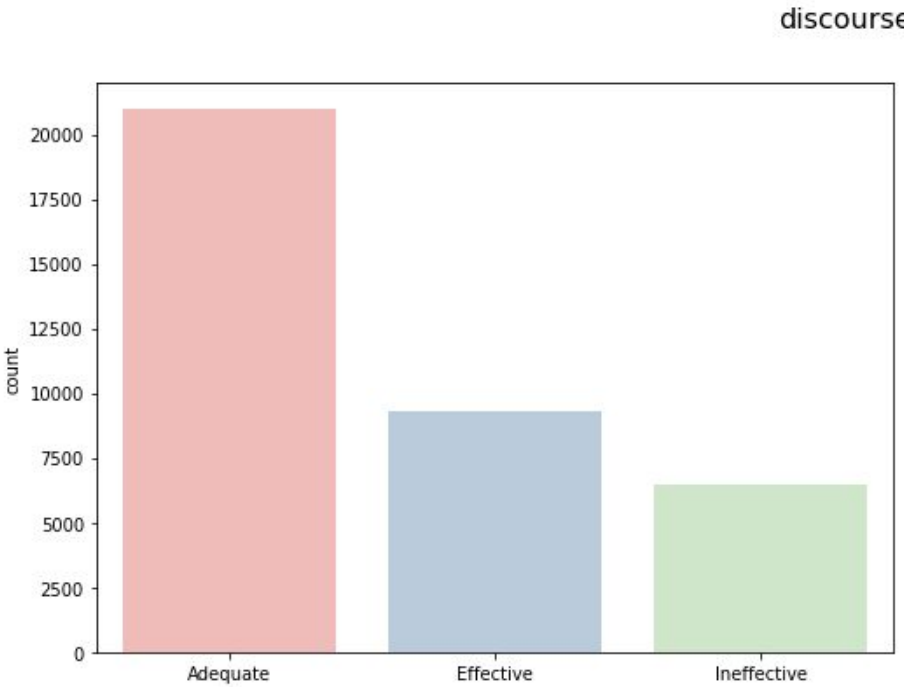
Discourse_id: 451b76cc4b59

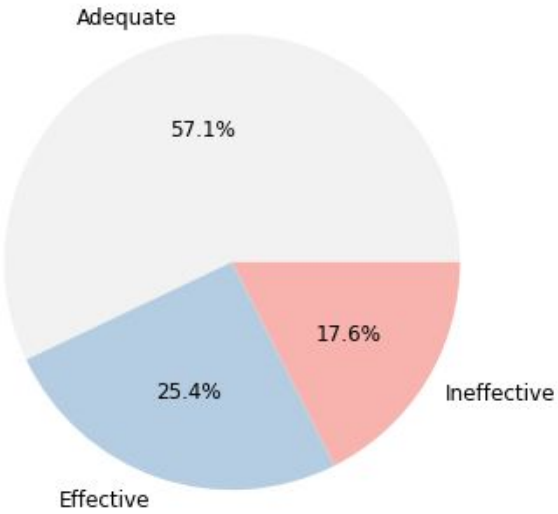
'After all, many students will greduate and keep going with their life without having any problem about jobs and they will be good economically. They also will give their childs a reason why they should go to school an d follow their dreams without letting people to\xa0punish their\xa0dreams. It also will help their family to support them without caring about what other people says, but everyone should conserve good things about ever ything. Also this students will have an illustrate part of his life, where they are going from success to suc cess keeping everything on the side but his dreams on the front. \xa0 \xa0 \xa0 '

After cleaning

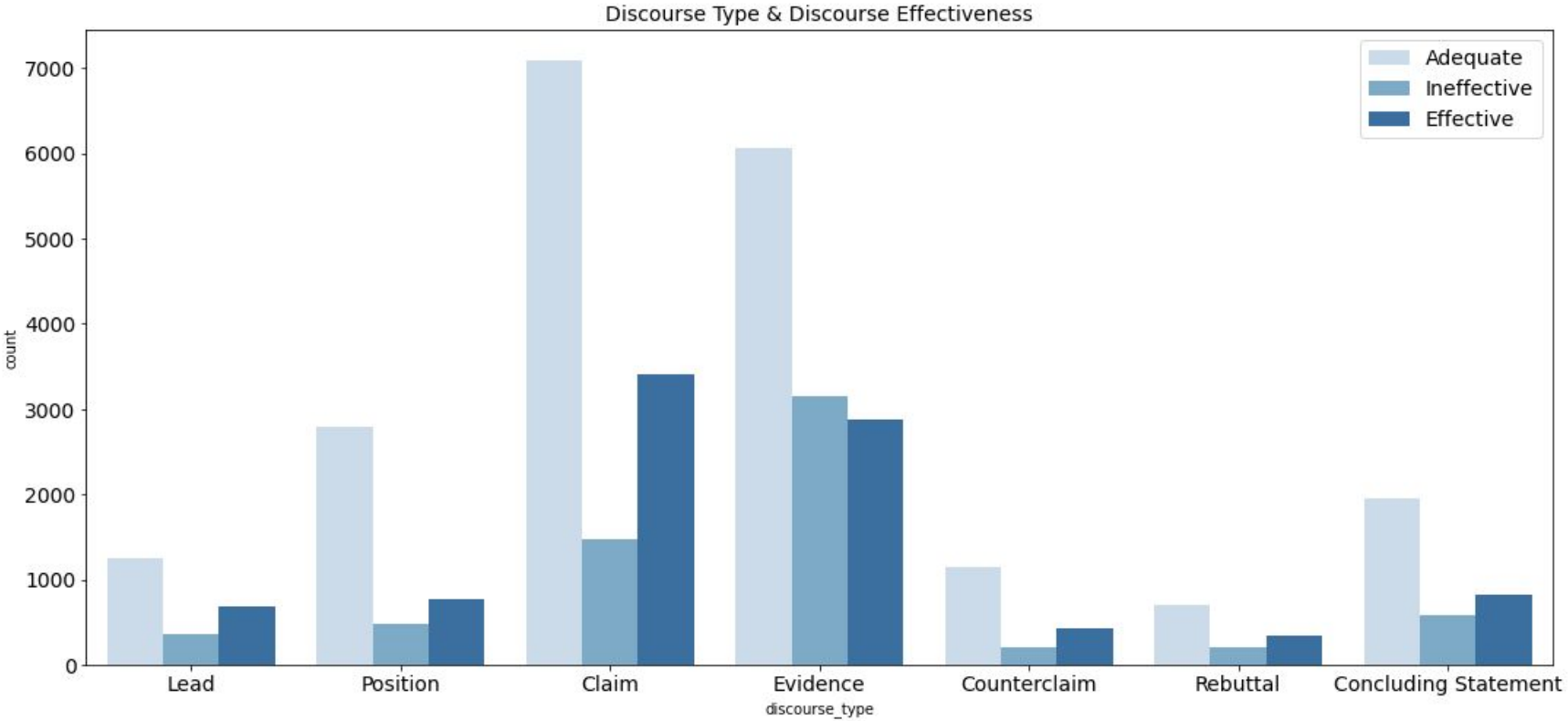
'After all, many students will greduate and keep going with their life without having any problem about jobs and they will be good economically. They also will give their childs a reason why they should go to school an d follow their dreams without letting people to punish their dreams. It also will help their family to suppor t them without caring about what other people says, but everyone should conserve good things about everythin g. Also this students will have an illustrate part of his life, where they are going from success to success keeping everything on the side but his dreams on the front. '

EDA - Target Class

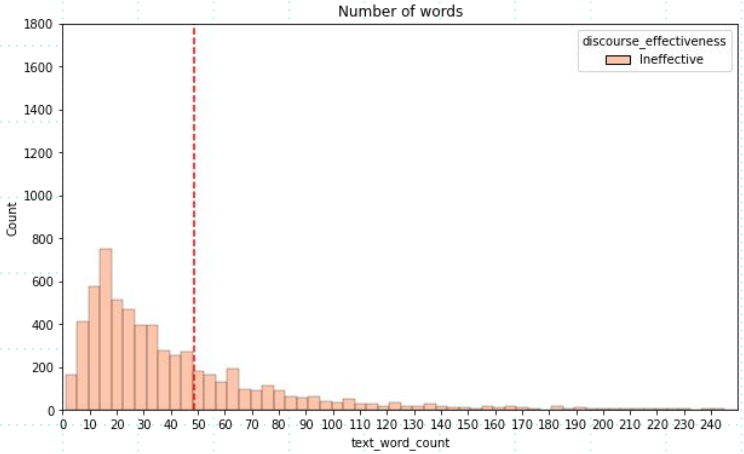


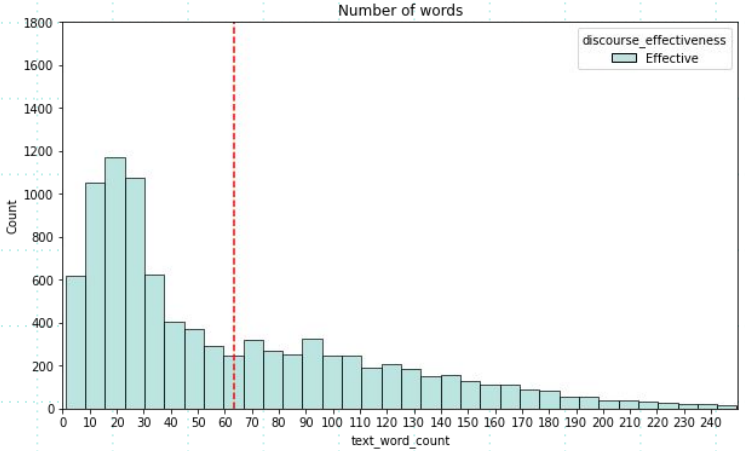


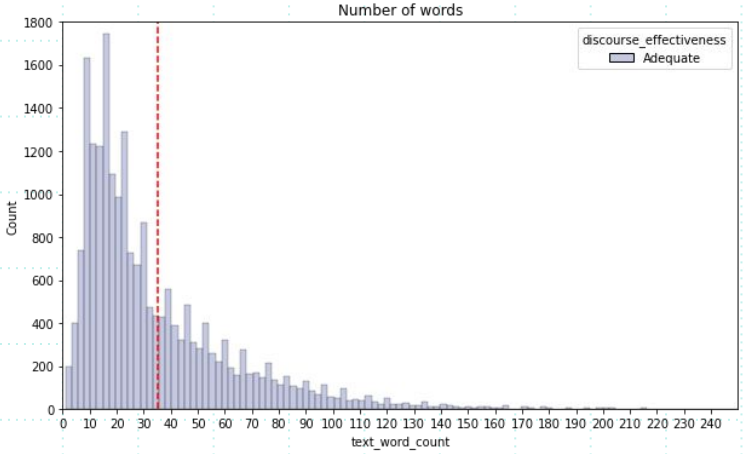
EDA - Discourse Type



EDA - Number of Words







Feature Engineering

1. Discourse Text

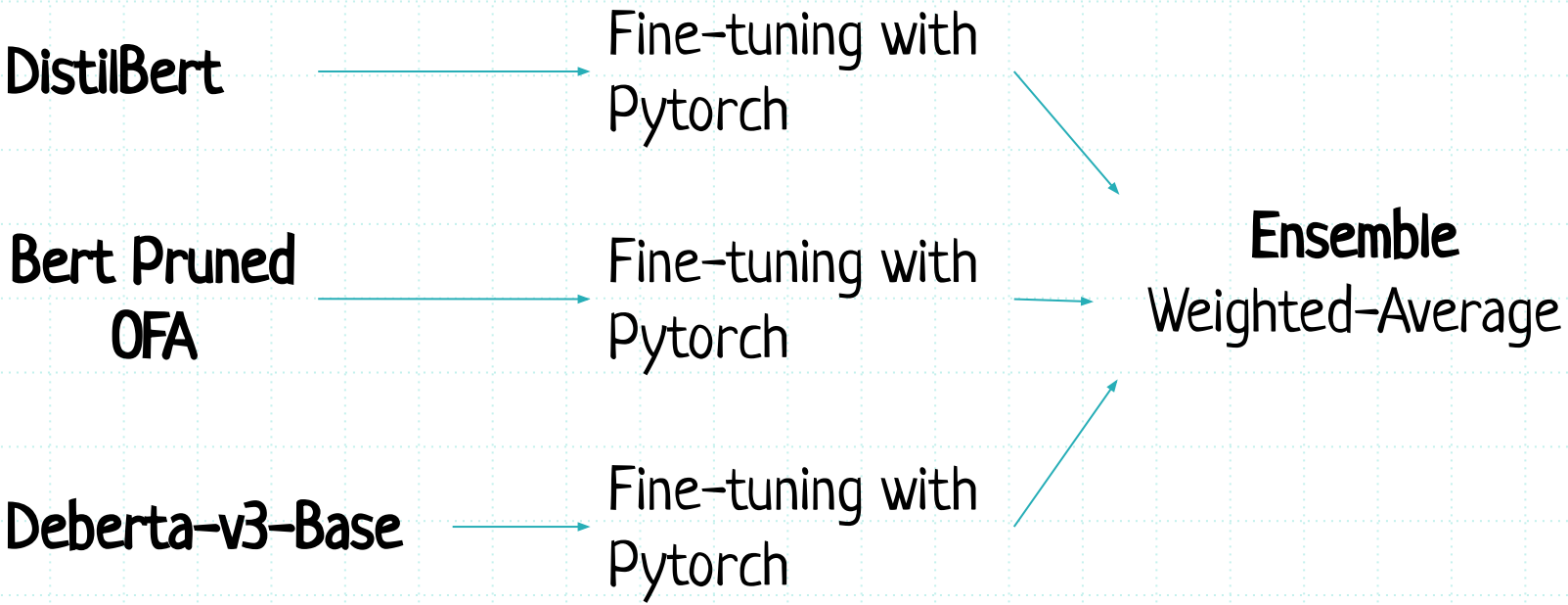
2. Discourse Type + Tokenizer Separator + Discourse Text

3. Discourse Type + Spacing + Discourse Text + Tokenizer Separator + Essay Text

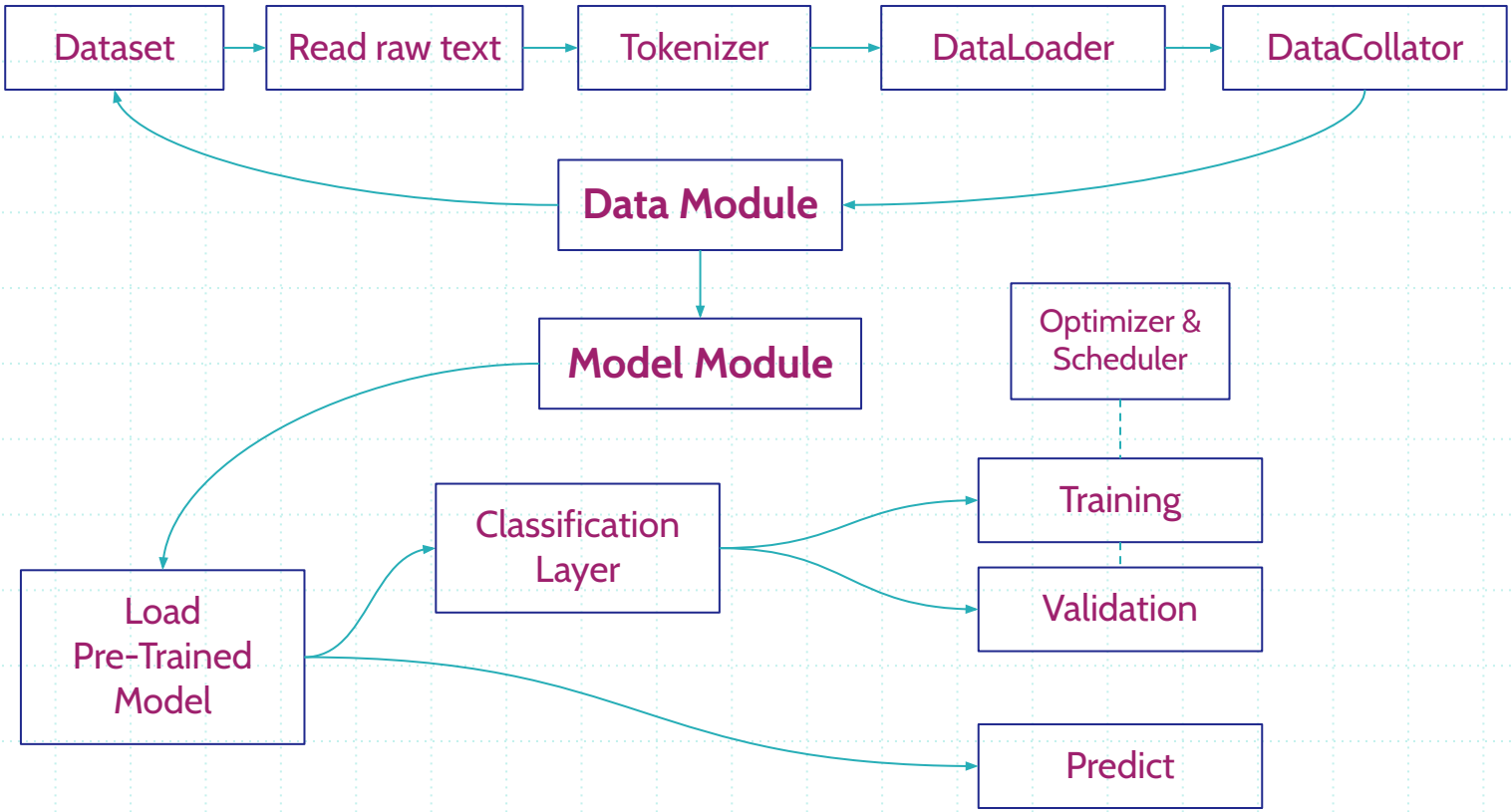
02

Modelling & Training

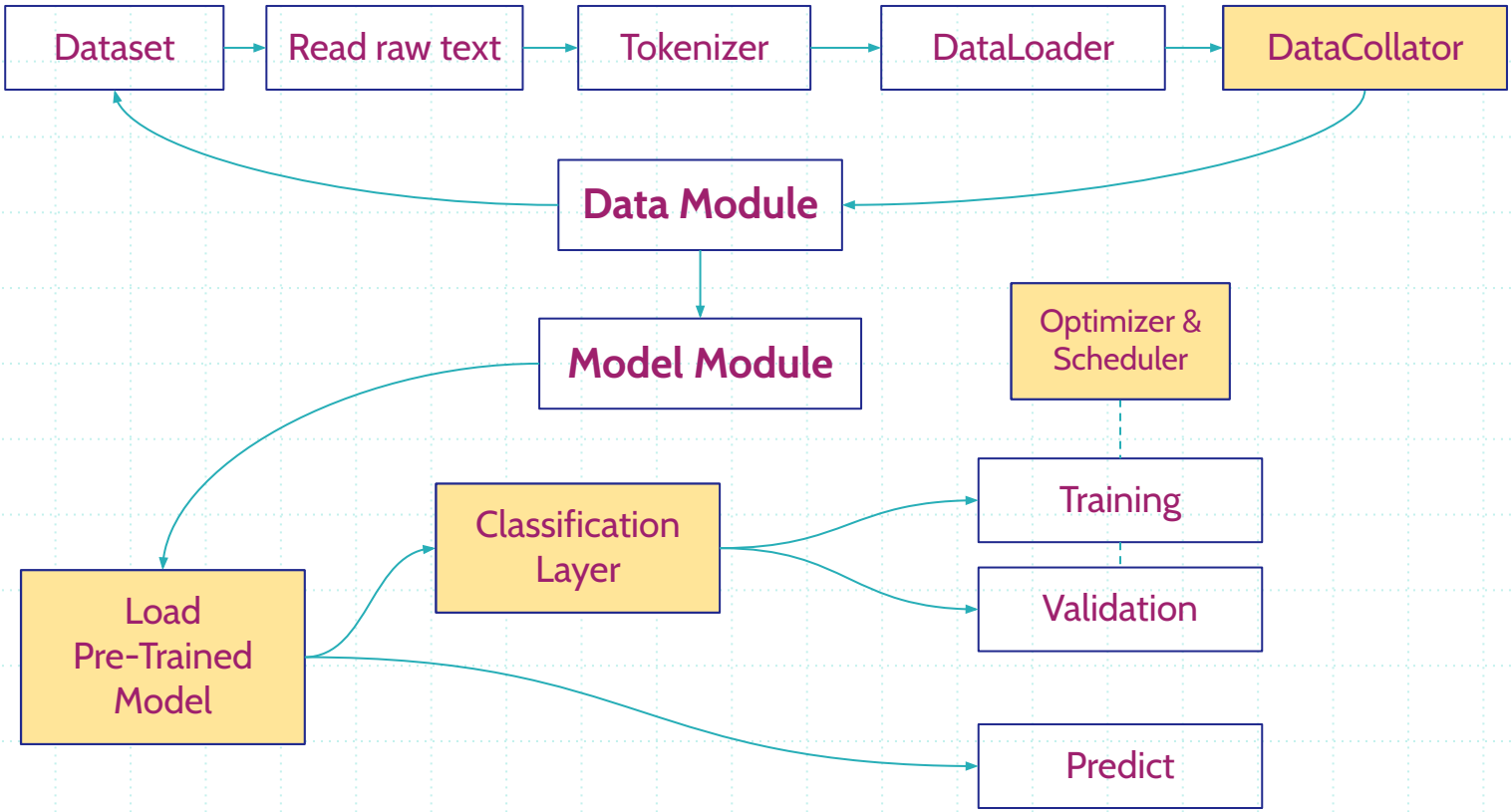
Approach - Fine Tuning PreTrained Models



Modelling Flow

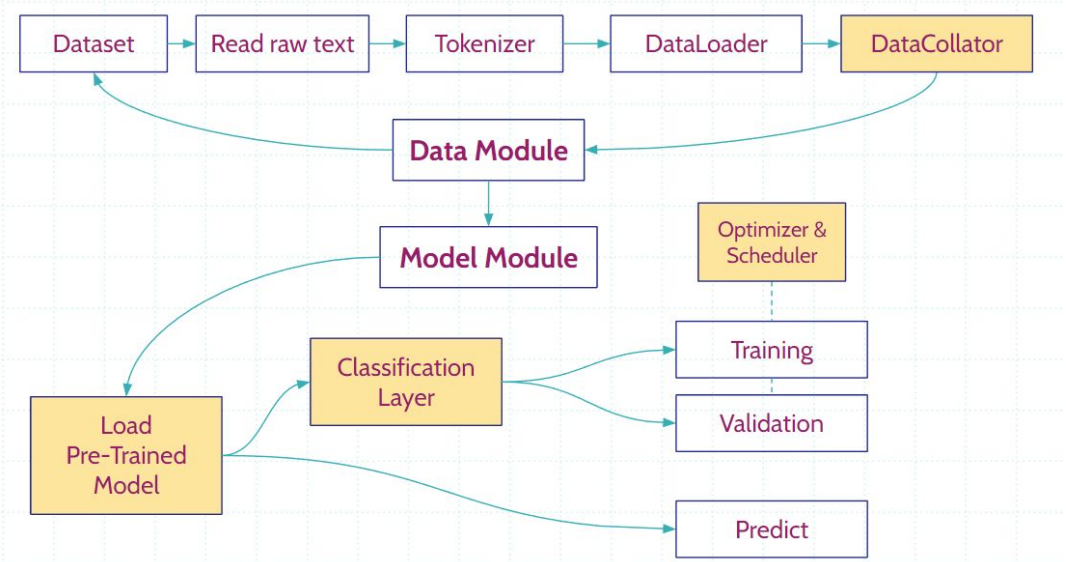


HyperParameter Tuning

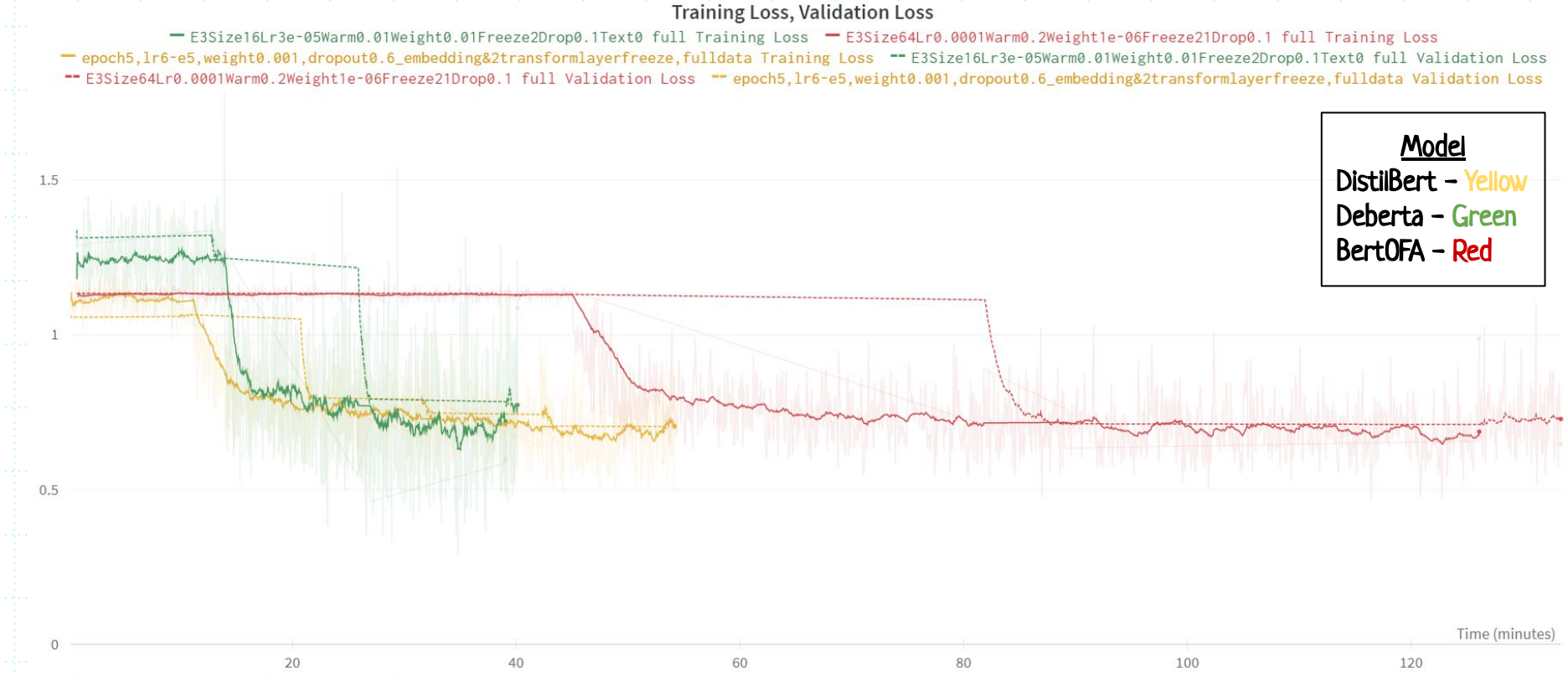


HyperParameter Tuning

1. Dynamic Padding
2. Layer Freezing
3. Average Pooling Layer
4. Additional Hidden Layer
5. Dropout
6. Learning Rate
7. Warmup



Training Time



03

Evaluation

Evaluation Metric

Differentiable

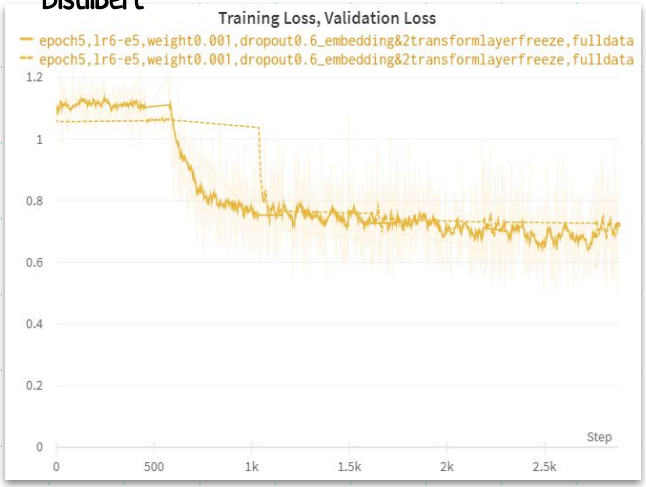
1. Cross Entropy Loss

- Multi-Class Classification

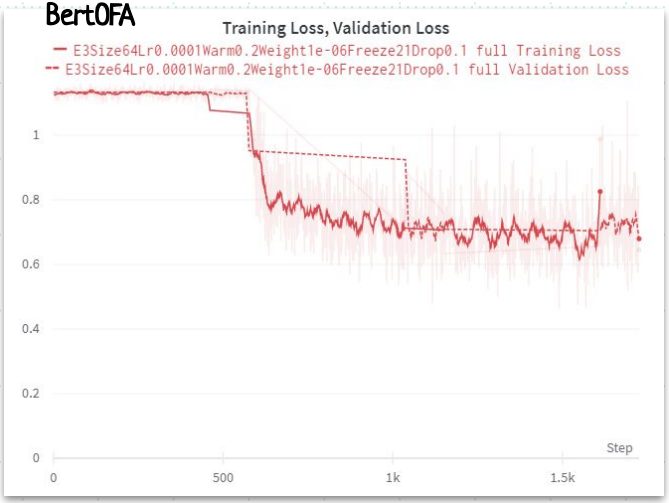
2. F1 – Score

- Imbalance dataset

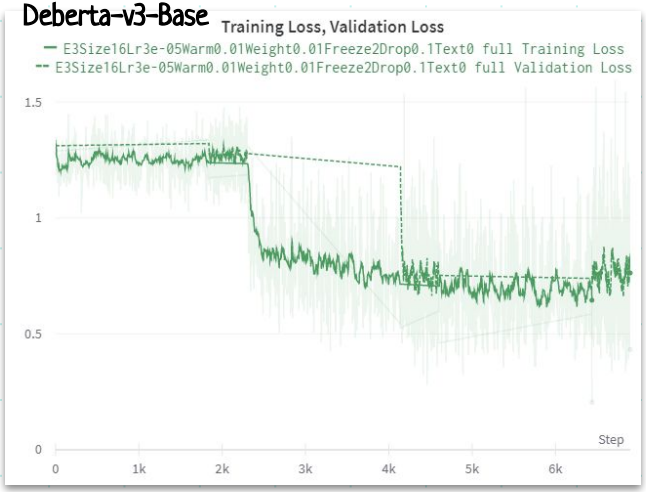
DistilBert



BertOFA



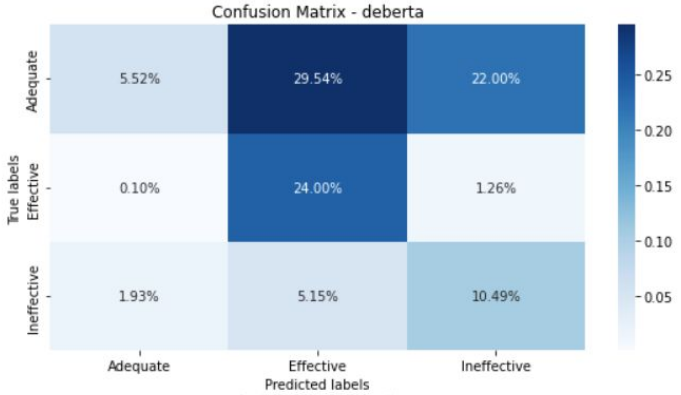
Deberta-v3-Base

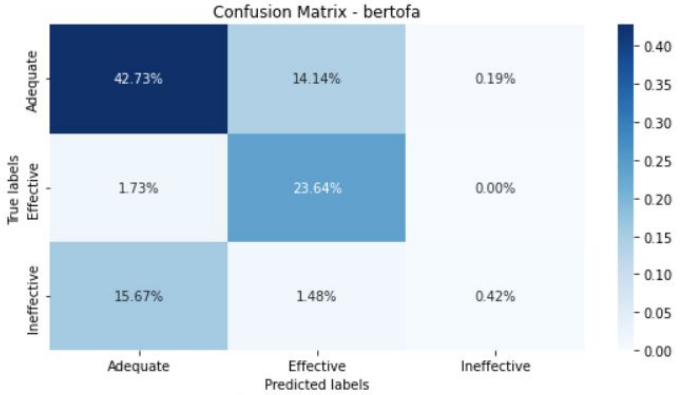


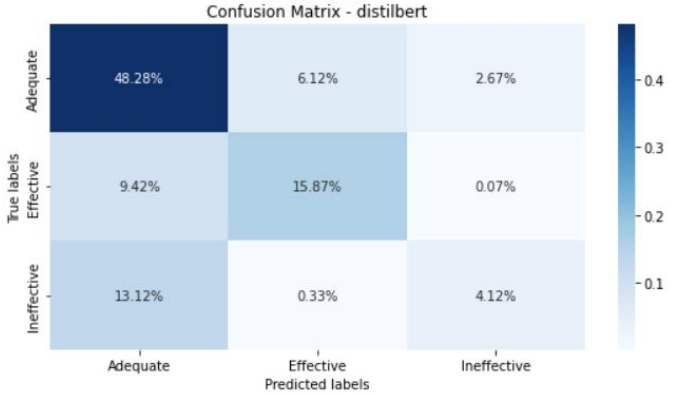
W&B

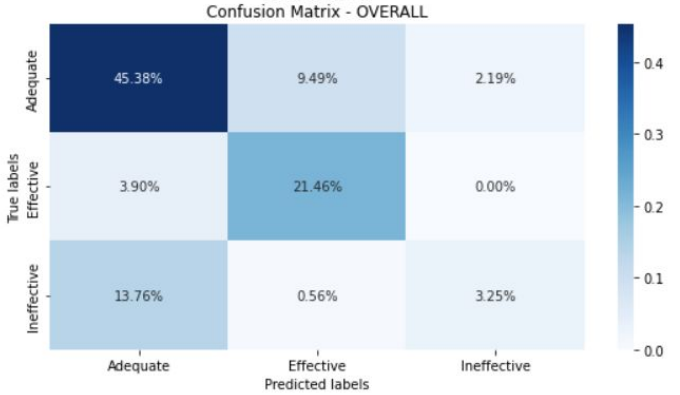
Interactive Link

Results









Weighted-Average

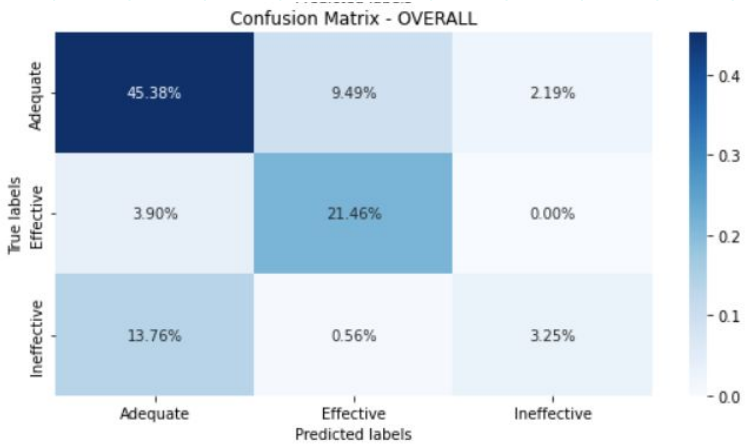
Results

F1 Score (Per Label)

label	deberta	distilbert	bertofa	overall
Adequate	0.170911	0.754998	0.729256	0.755661
Effective	0.571105	0.665716	0.731481	0.754663
Ineffective	0.408585	0.337416	0.046372	0.282506

- Ensemble technique:
  a. **Increases** f1 score
  b. **Reduces** log loss

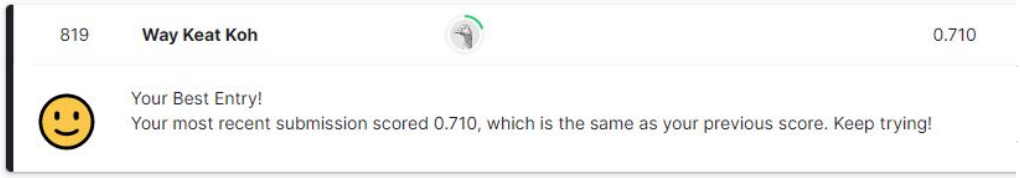
f1_weighted_score	log loss	model
0.314177	1.583759	deberta
0.658979	0.712088	distilbert
0.609831	0.728091	bertofa
0.672270	0.668112	OVERALL



Conclusion & Future Work

- Model performs well in classifying “Adequate” and “Effective” labels
- Model unable to classify “Ineffective” label effectively
- Ensemble was able to improve the score but not significantly

- Review hyperparameter
- Try batch normalization to reduce loss fluctuation during training
- Relook BertOFA model to ensure correct implementation
- Include weights in loss function to counter data imbalance



Thanks!

Does anyone have any questions?

waykeat@gmail.com

+65 91887276

CREDITS: This presentation template was created by **Slidesgo,** including icons by **Flaticon**, infographics & images by **Freepik**