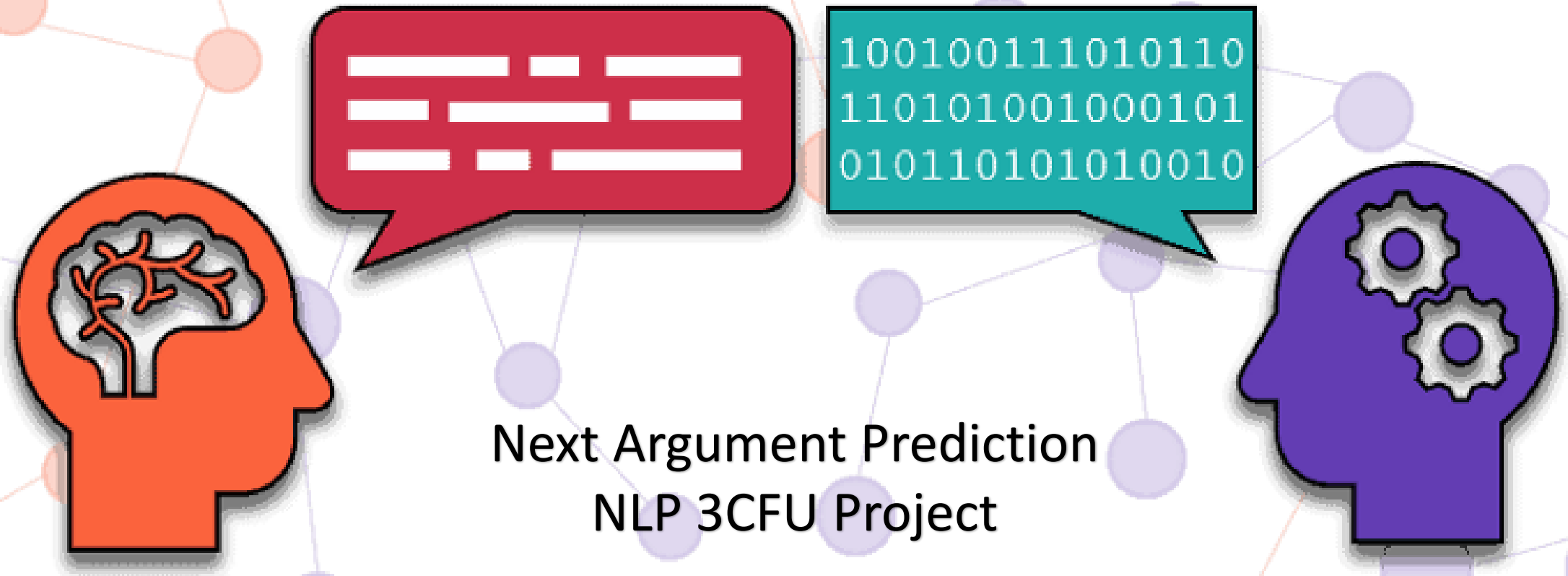
Natural Language Processing

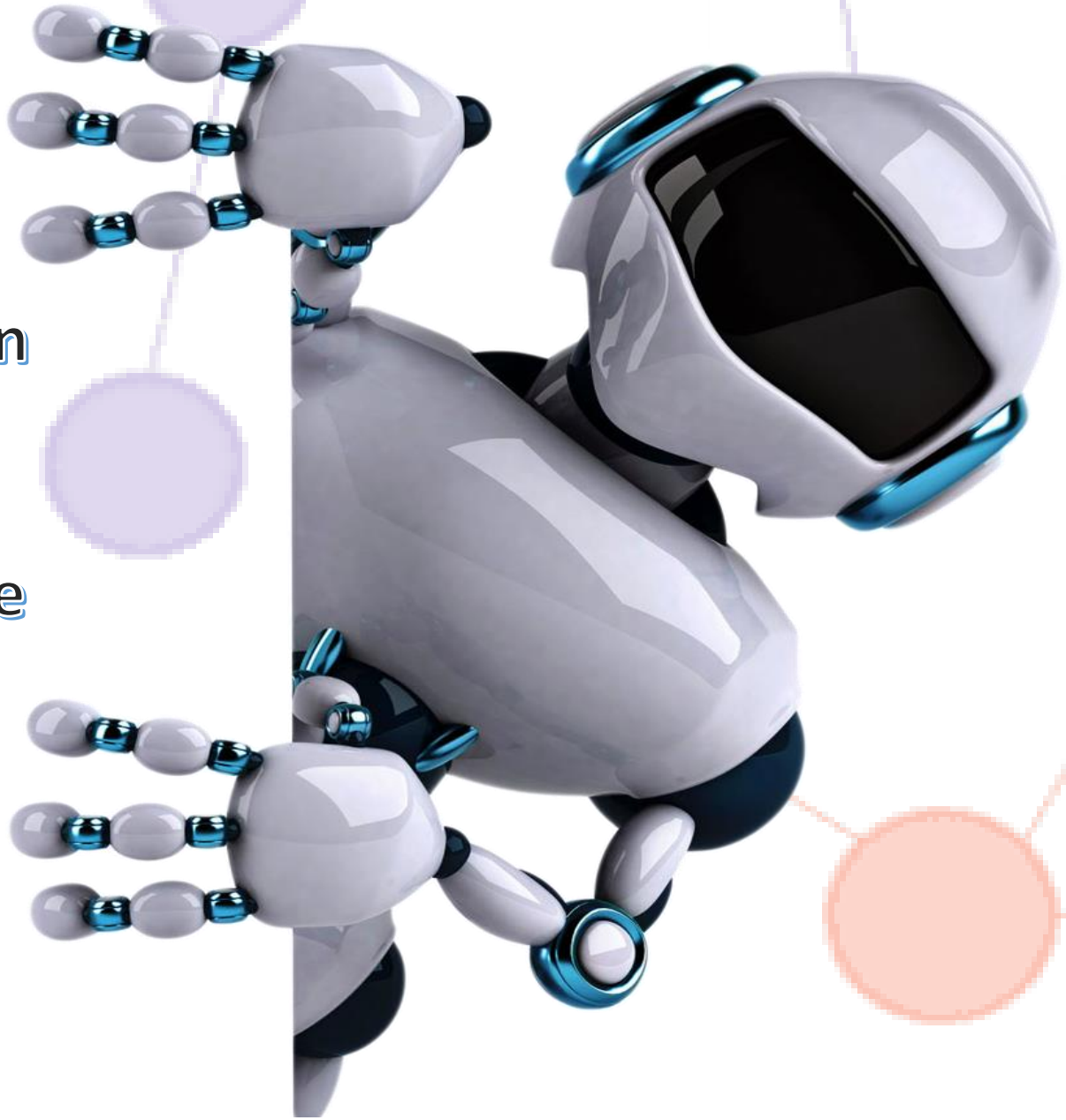


Next Argument Prediction
NLP 3CFU Project

Presented by Sami Osman

Outlines

1. Introduction
2. Objective
3. Dataset
4. Architecture
5. Training
6. Test



1. Introduction

Next sentence prediction is one of the many tasks in Natural language understanding which can be further seen as tasks like:

1. **The goal to identify whether the second sentence is entailment, contradiction, or neutral with respect to the first sentence.**
2. *The goal to determine whether two questions are semantically equal.*
3. *The task to determine whether the second sentence is the continuation of the first or not.*

Hypothesis

hypot: Eruptions can cause plants to die?

Text

sent1: eruptions emit lava.
sent2: eruptions produce ash clouds.
sent3: plants have green leaves.
sent4: producers will die without sunlight
sent5: ash blocks sunlight.

or

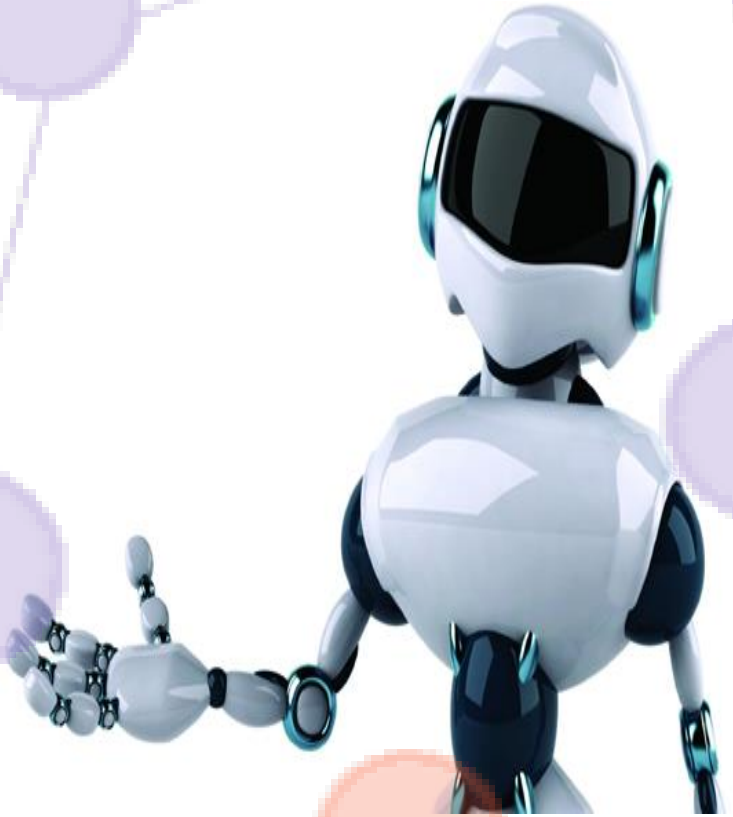
Corpus

2. Objective

In my work, I used a Transformer-based architecture, trained to solve a task that determine whether the second argument is the entailment of the first or not.

1. Using a dataset from Kialo.com website.
2. Finetuned a pre-trained BERT-based model.
3. Extend the network with LSTM/GRU.
4. Predict if argument1 precedes argument 2.

3. Dataset



Kialo Edu is the world's largest argument mapping and debate site.

Raw data: The downloaded files have a tree like data structure where people write their argument based on the previous argument or related to the discussion topic.

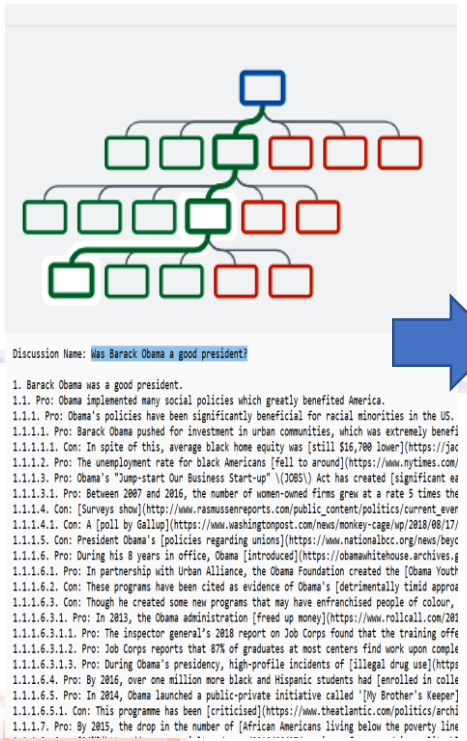
First transformation: This dataset is consisting of all possible details of the file downloaded from Kialo website.

Second transformation: The dataset is consisting of a special format which is consistent to the task on hand (NSP). At this point the dataset has the structure [index, label, argument1, argument2].

The text preprocessing includes make text lowercase, remove text in square brackets, remove links, remove punctuation and removing words containing numbers.

Dataset ...

Downloaded around 300 files from kialo.com site and made ≈80,000 trainset from 190 files, ≈20,000 validation set from 70 files and 10,000 test set from 40 files. The files are stored in separate folders and the code goes through each folder to create train, validation and test sets.



Discussion Name: Was Barack Obama a good president?

1. Barack Obama was a good president.
1.1. Pro: Obama implemented many social policies which greatly benefited America.
1.1.1. Pro: Obama's policies have been significantly beneficial for racial minorities in the US.
1.1.1.1. Pro: Barack Obama pushed for investment in urban communities, which was extremely benefi
1.1.1.1.1. Con: In spite of this, average black home equity was [still $16,700 lower](https://jac
1.1.1.2. Pro: The unemployment rate for black Americans [fell to around](https://www.nytimes.com/
1.1.1.3. Pro: Obama's "Jump-start Our Business Start-up" \(JOBS\) Act has created [significant ea
1.1.1.3.1. Pro: Between 2007 and 2016, the number of women-owned firms grew at a rate 5 times the
1.1.1.4. Con: [Surveys show](http://www.rasmussenreports.com/public_content/politics/current_ever
1.1.1.4.1. Con: A [poll by Gallup](https://www.washingtonpost.com/news/monkey-cage/wp/2018/08/17/
1.1.1.5. Con: President Obama's [policies regarding unions](https://www.nationalbcc.org/news/beyc
1.1.1.6. Pro: During his 8 years in office, Obama [introduced](https://obamawhitehouse.archives.g
1.1.1.6.1. Pro: In partnership with Urban Alliance, the Obama Foundation created the [Obama Youth
1.1.1.6.2. Con: These programs have been cited as evidence of Obama's [detrimentally timid approa
1.1.1.6.3. Con: Though he created some new programs that may have enfranchised people of colour,
1.1.1.6.3.1. Pro: In 2013, the Obama administration [freed up money](https://www.rollcall.com/201
1.1.1.6.3.1.1. Pro: The inspector general's 2018 report on Job Corps found that the training offe
1.1.1.6.3.1.2. Pro: Job Corps reports that 87% of graduates at most centers find work upon comple
1.1.1.6.3.1.3. Pro: During Obama's presidency, high-profile incidents of [illegal drug use](https
1.1.1.6.4. Pro: By 2016, over one million more black and Hispanic students had [enrolled in colle
1.1.1.6.5. Pro: In 2014, Obama launched a public-private initiative called "[My Brother's Keeper]
1.1.1.6.5.1. Con: This programme has been [criticised](https://www.theatlantic.com/politics/archi
1.1.1.7. Pro: By 2015, the drop in the number of [African Americans living below the poverty line

	index	title	position	opinion	argument
2	0	Was Barack Obama a good	[1]	Farg	Obama was a good president.
3	1	Was Barack Obama a good	[1, 1]	Pro:	Obama implemented many social policies which greatly benefited America.
4	2	Was Barack Obama a good	[1, 1, 1]	Pro:	Obama's policies have been significantly beneficial for racial minorities in the US.
5	3	Was Barack Obama a good	[1, 1, 1, 1]	Pro:	Barack Obama pushed for investment in urban communities, which was extremely beneficial for the [growth of black wealth](https://www.aljazeera.com/programm
6	4	Was Barack Obama a good	[1, 1, 1, 1,	Con:	In spite of this, average black home equity was [still $16,700 lower](https://jacobinmag.com/2017/12/obama-foreclosure-crisis-wealth-inequality) in 2016 than it wa
7	5	Was Barack Obama a good	[1, 1, 1, 2]	Pro:	The unemployment rate for black Americans [fell to around](https://www.nytimes.com/2018/08/14/us/politics/fact-check-trump-jobs-black-americans.html) 7.8% d
8	6	Was Barack Obama a good	[1, 1, 1, 3]	Pro:	Obama's "Jump-start Our Business Start-up" \(JOBS\) Act has created [significant ease](https://www.blackenterprise.com/equity-crowdfunding-black-investors-entre
9	7	Was Barack Obama a good	[1, 1, 1, 3,	Pro:	Between 2007 and 2016, the number of women-owned firms grew at a rate 5 times the national average, more than doubling the number of firms owned by [Africa
10	8	Was Barack Obama a good	[1, 1, 1, 4]	Con:	[Surveys show](http://www.rasmussenreports.com/public_content/politics/current_events/social_issues/voters_say_trump_better_for_blacks_than_obama_but_m
11	9	Was Barack Obama a good	[1, 1, 1, 4,	Con:	A [poll by Gallup](https://www.washingtonpost.com/news/monkey-cage/wp/2018/08/17/no-one-third-of-african-americans-dont-support-trump-not-even-close/) sl
12	10	Was Barack Obama a good	[1, 1, 1, 5]	Con:	President Obama's [policies regarding unions](https://www.nationalbcc.org/news/beyond-the-rhetoric/1176-president-obama-is-selling-out-blacks-for-union-favor)
13	11	Was Barack Obama a good	[1, 1, 1, 6]	Pro:	During his 8 years in office, Obama [introduced](https://obamawhitehouse.archives.gov/the-press-office/2016/10/14/progress-african-american-community-during-
14	12	Was Barack Obama a good	[1, 1, 1, 6,	Pro:	In partnership with Urban Alliance, the Obama Foundation created the [Obama Youth Jobs Corps](https://theurbanalliance.org/our-programs/obama-y
15	13	Was Barack Obama a good	[1, 1, 1, 6,	Con:	These programs have been cited as evidence of Obama's [detrimentally timid approach](https://www.theatlantic.com/politics/archive/2016/12/how-ba
16	14	Was Barack Obama a good	[1, 1, 1, 6,	Con:	Though he created some new programs that may have enfranchised people of colour, Obama also allowed the continuation of established ones that had a history
17	15	Was Barack Obama a good	[1, 1, 1, 6,	Pro:	In 2013, the Obama administration [freed up money](https://www.rollcall.com/2013/05/15/job-corps-hits-turning-point-under-congress-watch/) from other employ
18	16	Was Barack Obama a good	[1, 1, 1, 6,	Pro:	The inspector general's 2018 report on Job Corps found that the training offered [no discernible long-term benefit](https://www.nytimes.com/2018/08/26/us/politi
19	17	Was Barack Obama a good	[1, 1, 1, 6,	Pro:	Job Corps reports that 87% of graduates at most centers find work upon completion of the program; however, many of these jobs are in the [same low-wage, low-s
20	18	Was Barack Obama a good	[1, 1, 1, 6,	Pro:	During Obama's presidency, high-profile incidents of [illegal drug use](https://youthtoday.org/2009/10/how-not-to-handle-sex-drugs-and-violence-at-job-corps/), [v
21	19	Was Barack Obama a good	[1, 1, 1, 6,	Pro:	By 2016, over one million more black and Hispanic students had [enrolled in college](https://obamawhitehouse.archives.gov/the-press-office/2016/10/14/progress-
22	20	Was Barack Obama a good	[1, 1, 1, 6,	Pro:	In 2014, Obama launched a public-private initiative called "[My Brother's Keeper](https://theundefeated.com/features/my-brothers-keeper-and-obamas-mission-to-
23	21	Was Barack Obama a good	[1, 1, 1, 6,	Con:	This programme has been [criticised](https://www.theatlantic.com/politics/archive/2016/12/how-barack-obama-failed-black-americans/511358/) for its misguided f
24	22	Was Barack Obama a good	[1, 1, 1, 7]	Pro:	By 2015, the drop in the number of [African Americans living below the poverty line](https://obamawhitehouse.archives.gov/the-press-office/2016/10/14/progress-
25	23	Was Barack Obama a good	[1, 1, 1, 8]	Con:	[34%](https://www.pewsocialtrends.org/2016/06/27/on-views-of-race-and-inequality-blacks-and-whites-are-worlds-apart/) of black Americans think Obama has trie
26	24	Was Barack Obama a good	[1, 1, 2]	Pro:	Obama implemented the [Deferred Action for Childhood Arrivals Act](https://www.dhs.gov/deferred-action-childhood-arrivals-daca) \(DACA\), which benefited unc
27	25	Was Barack Obama a good	[1, 1, 2, 1]	Pro:	The DACA afforded undocumented immigrants [access](https://www.vox.com/2017/9/2/16244380/daca-benefits-trump-undocumented-immigrants-jobs) to greater
28	26	Was Barack Obama a good	[1, 1, 2, 2]	Pro:	The DACA [shielded](https://www.vox.com/2017/9/2/16244380/daca-benefits-trump-undocumented-immigrants-jobs) undocumented immigrants from deportatio
29	27	Was Barack Obama a good	[1, 1, 2, 3]	Con:	While many Obama supporters criticise Trump's tough stance on immigration, Obama deported [significantly more](https://www.wsj.com/articles/why-trump-has-d
30	28	Was Barack Obama a good	[1, 1, 2, 3,	Con:	A numbers-by-numbers contrast of immigrants deported arguably does not provide an accurate picture of Obama's stance on immigrant rights as compared with T

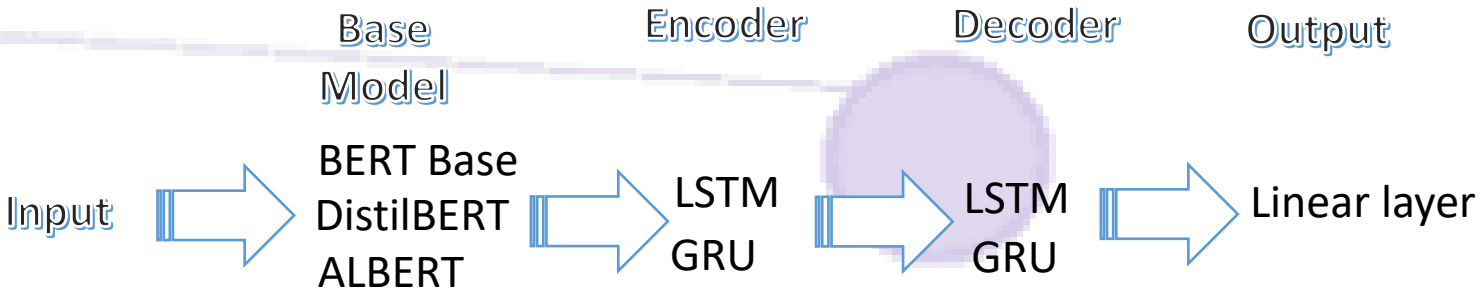
	index	label	position1	position2	argument1	argument2
0	0	[1]	[1, 1]		Obama was a good presid	Obama implemented many social policies which greatly benefited America.
1	1	[1, 1]	[1]		Obama implemented man	Obama was a good president.
2	0	[1, 1]	[1, 1, 1]		Obama implemented man	Obama's policies have been significantly beneficial for racial minorities in the US.
3	1	[1, 1]	[1, 2, 2, 3,		Obama implemented man	Despite an increase in numbers, by the end of Obama's presidency the Afghan forces still [lacked](https://www.politifact.com/truth-o-meter/promises/obameter/promise/165/train-and
4	0	[1, 1, 1]	[1, 1, 1, 1]		Obama's policies have bee	Barack Obama pushed for investment in urban communities, which was extremely beneficial for the [growth of black wealth](https://www.aljazeera.com/programmes/upfront/2018/10
5	1	[1, 1, 1, 1]	[1, 1, 1]		Barack Obama pushed for	Obama's policies have been significantly beneficial for racial minorities in the US.
6	0	[1, 1, 1, 1]	[1, 1, 1, 1,		Barack Obama pushed for	In spite of this, average black home equity was [still $16,700 lower](https://jacobinmag.com/2017/12/obama-foreclosure-crisis-wealth-inequality) in 2016 than it was before the Obama
7	1	[1, 1, 1, 1]	[1, 1, 2, 3,		Barack Obama pushed for	[Statistics by](https://www.axios.com/immigration-ice-deportation-trump-obama-a72a0a44-540d-46bc-a671-cd65c772f4b1.html) the Department of Homeland Security show that Oba
8	0	[1, 1, 1]	[1, 1, 1, 2]		Obama's policies have bee	The unemployment rate for black Americans [fell to around](https://www.nytimes.com/2018/08/14/us/politics/fact-check-trump-jobs-black-americans.html) 7.8% during the Obama ad
9	1	[1, 1, 1, 2]	[1, 1, 1]		The unemployment rate fc	Obama's policies have been significantly beneficial for racial minorities in the US.
10	0	[1, 1, 1]	[1, 1, 1, 3]		Obama's policies have bee	Obama's "Jump-start Our Business Start-up" \(JOBS\) Act has created [significant ease](https://www.blackenterprise.com/equity-crowdfunding-black-investors-entrepreneurs/) for mino
11	1	[1, 1, 1]	[1, 5, 3, 1,		Obama's policies have bee	Some blame the [removal](https://fortune.com/2017/08/04/dodd-frank-choice-act/) of restrictions on multistate banking, and not the Dodd-Frank Act, for the decline of small banks, le
12	0	[1, 1, 1, 3]	[1, 1, 1, 3,		Obama's "Jump-start Our	Between 2007 and 2016, the number of women-owned firms grew at a rate 5 times the national average, more than doubling the number of firms owned by [African American women a
13	1	[1, 1, 1, 3,	[1, 1, 1, 3]		Between 2007 and 2016, t	Obama's "Jump-start Our Business Start-up" \(JOBS\) Act has created [significant ease](https://www.blackenterprise.com/equity-crowdfunding-black-investors-entrepreneurs/) for mino
14	0	[1, 1, 1]	[1, 1, 1, 4]		Obama's policies have bee	[Surveys show](http://www.rasmussenreports.com/public_content/politics/current_events/social_issues/voters_say_trump_better_for_blacks_than_obama_but_more_is_needed) tha
15	1	[1, 1, 1]	[1, 6, 8, 2,		Obama's policies have bee	According to many reports, Obama felt that Hillary Clinton's presidency would further his own policy agenda, and hence solidify and advance his legacy.
16	0	[1, 1, 1, 4]	[1, 1, 1, 4,		[Surveys show](http://ww	A [poll by Gallup](https://www.washingtonpost.com/news/monkey-cage/wp/2018/08/17/no-one-third-of-african-americans-dont-support-trump-not-even-close/) showed that these a
17	1	[1, 1, 1, 4,	[1, 1, 1, 4]		A [poll by Gallup](https://v	[Surveys show](http://www.rasmussenreports.com/public_content/politics/current_events/social_issues/voters_say_trump_better_for_blacks_than_obama_but_more_is_needed) tha
18	0	[1, 1, 1]	[1, 1, 1, 5]		Obama's policies have bee	President Obama's [policies regarding unions](https://www.nationalbcc.org/news/beyond-the-rhetoric/1176-president-obama-is-selling-out-blacks-for-union-favor) have been extremel
19	1	[1, 1, 1]	[1, 6, 8, 3,		Obama's policies have bee	Obama's election as the first black president may have given his presidency [fake-progressive clothing](https://www.counterpunch.org/2012/03/08/is-barack-obama-a-fit-role-model-fo
20	0	[1, 1, 1]	[1, 1, 1, 6]		Obama's policies have bee	During his 8 years in office, Obama [introduced](https://obamawhitehouse.archives.gov/the-press-office/2016/10/14/progress-african-american-community-during-obama-administrati
21	1	[1, 1, 1, 6]	[1, 1, 1]		During his 8 years in office	Obama's policies have been significantly beneficial for racial minorities in the US.
22	0	[1, 1, 1, 6]	[1, 1, 1, 6,		During his 8 years in office	In partnership with Urban Alliance, the Obama Foundation created the [Obama Youth Jobs Corps](https://theurbanalliance.org/our-programs/obama-youth-jobs-corps/) \(OYJC\) to incr
23	1	[1, 1, 1, 6]	[1, 4, 5]		During his 8 years in office	The Obama administration created strains in the relationships between the US and its allies at the time.
24	0	[1, 1, 1, 6]	[1, 1, 1, 6,		During his 8 years in office	These programs have been cited as evidence of Obama's [detrimentally timid approach](https://www.theatlantic.com/politics/archive/2016/12/how-barack-obama-failed-black-americ
25	1	[1, 1, 1, 6,	[1, 1, 1, 6]		During his 8 years in office, Obama [introduced](https://obamawhitehouse.archives.gov/the-press-office/2016/10/14/progress-african-american-community-during-obama-administrati	These programs have been
26	0	[1, 1, 1, 6]	[1, 1, 1, 6,		During his 8 years in office	Though he created some new programs that may have enfranchised people of colour, Obama also allowed the continuation of established ones that had a history of poor success.
27	1	[1, 1, 1, 6]	[1, 4, 4]		During his 8 years in office	The Obama administration has repeatedly failed in the Middle East.
28	0	[1, 1, 1, 6,	[1, 1, 1, 6,		Though he created some n	In 2013, the Obama administration [freed up money](https://www.rollcall.com/2013/05/15/job-corps-hits-turning-point-under-congress-watch/) from other employment training progra
29	1	[1, 1, 1, 6,	[1, 1, 1, 6,		In 2013, the Obama admin	Though he created some new programs that may have enfranchised people of colour, Obama also allowed the continuation of established ones that had a history of poor success.
30	0	[1, 1, 1, 6,	[1, 1, 1, 6,		In 2013, the Obama admin	The inspector general's 2018 report on Job Corps found that the training offered [no discernible long-term benefit](https://www.nytimes.com/2018/08/26/us/politics/job-corps-training

4. Architecture

Three base architectures all of which are mainly based on pre-trained Transformers.

The first experiment is fine-tuning Bert that is linked to a linear binary classification network.

In the following phases I modified this baseline model, by adding or modifying the architecture on top of the base model (i.e., LSTM and GRU encoder-decoder)



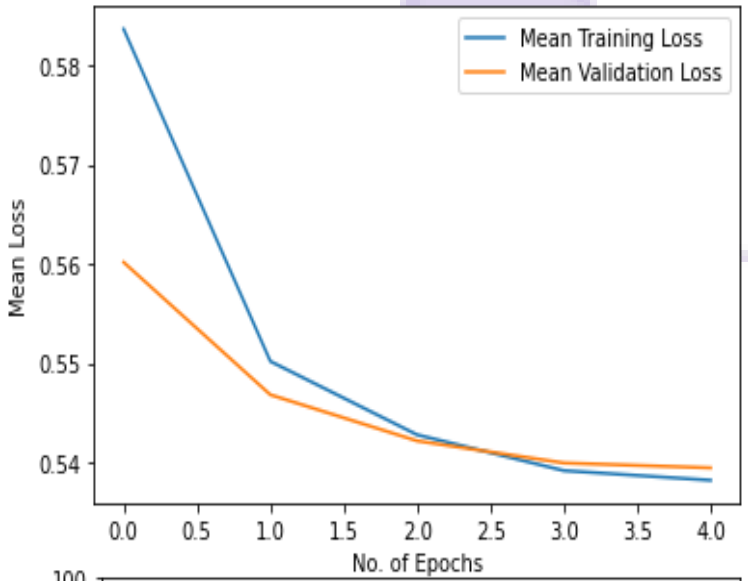
Hyper parameters:

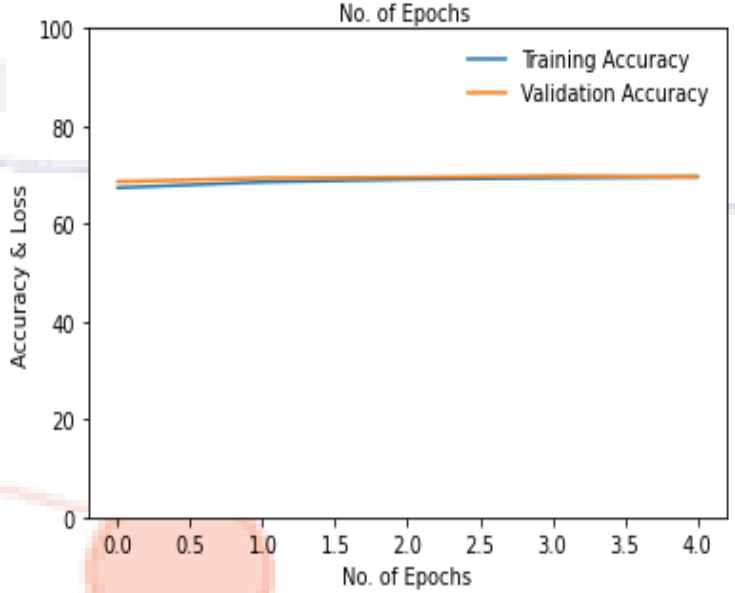
I tested each model by running a 5 training epoch, using a learning rate of 2e-5 and a dropout rate of 0.2. I used Cross Entropy loss and AdamW optimizer.

5. Training

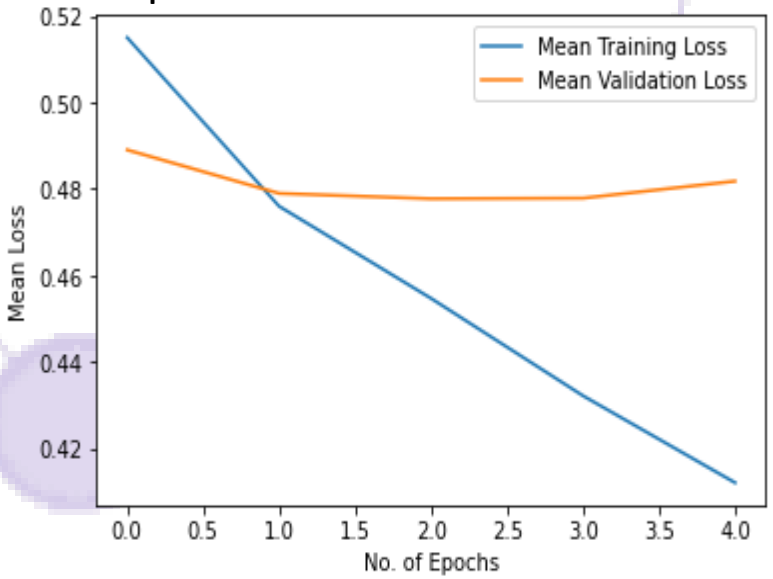
Albert – V[0,1,2]

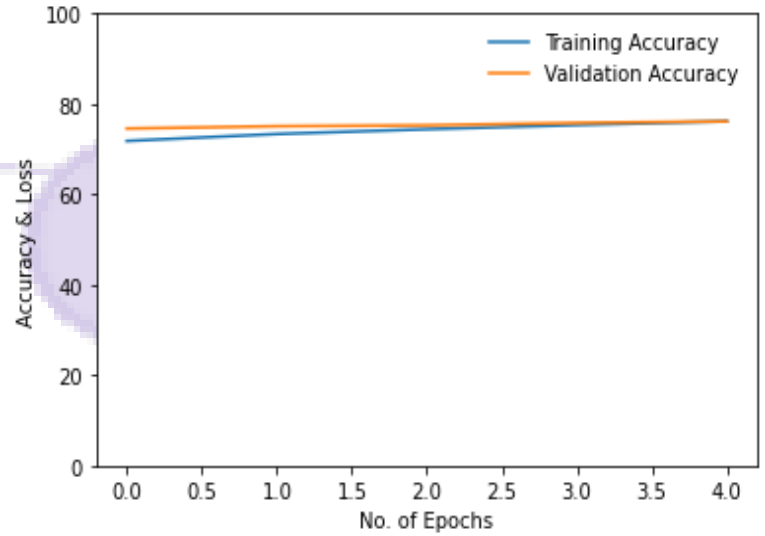
Experiment Phase 1



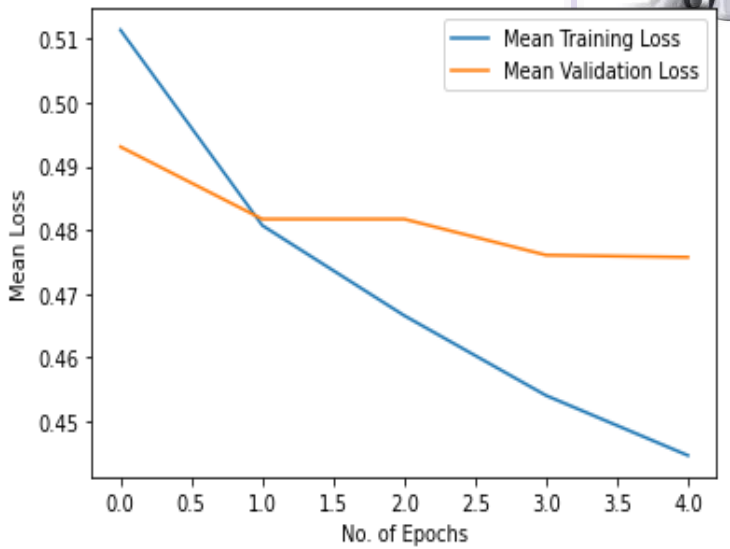


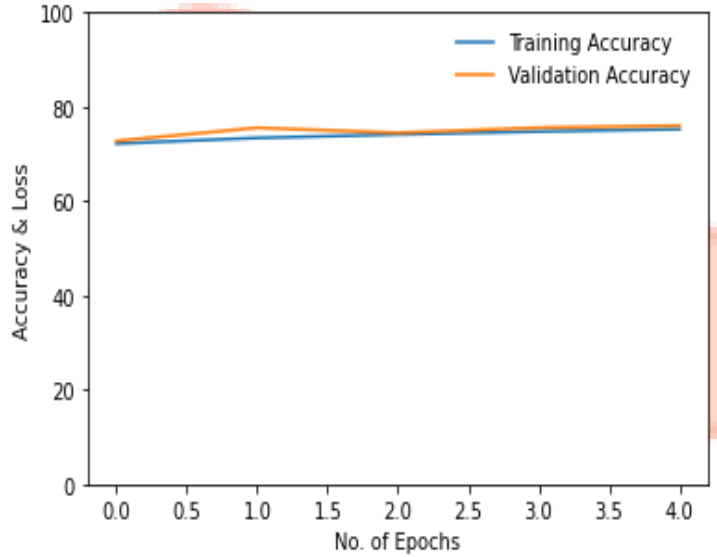
Experiment Phase 2





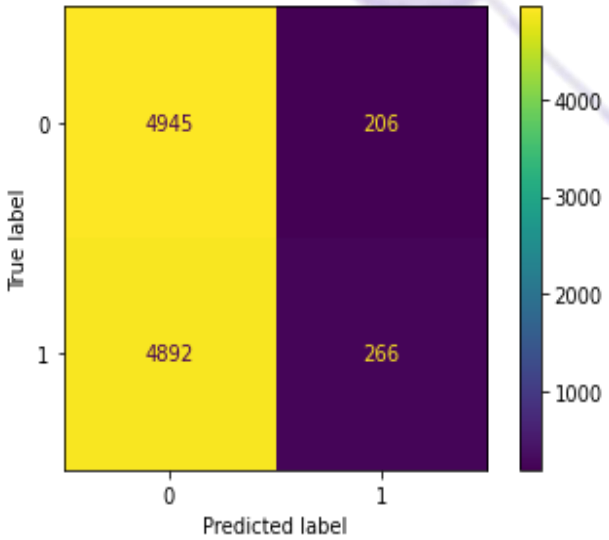
Experiment Phase 3



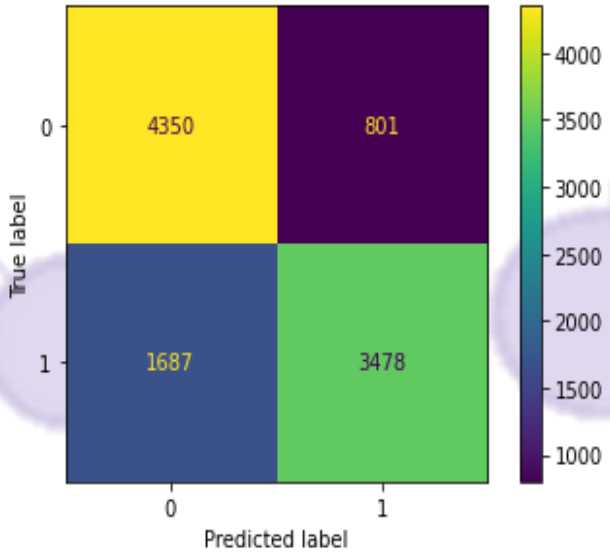


6. Test

Transpose + DistilBert Confusion Matrix



No Transpose + DistilBert Confusion Matrix

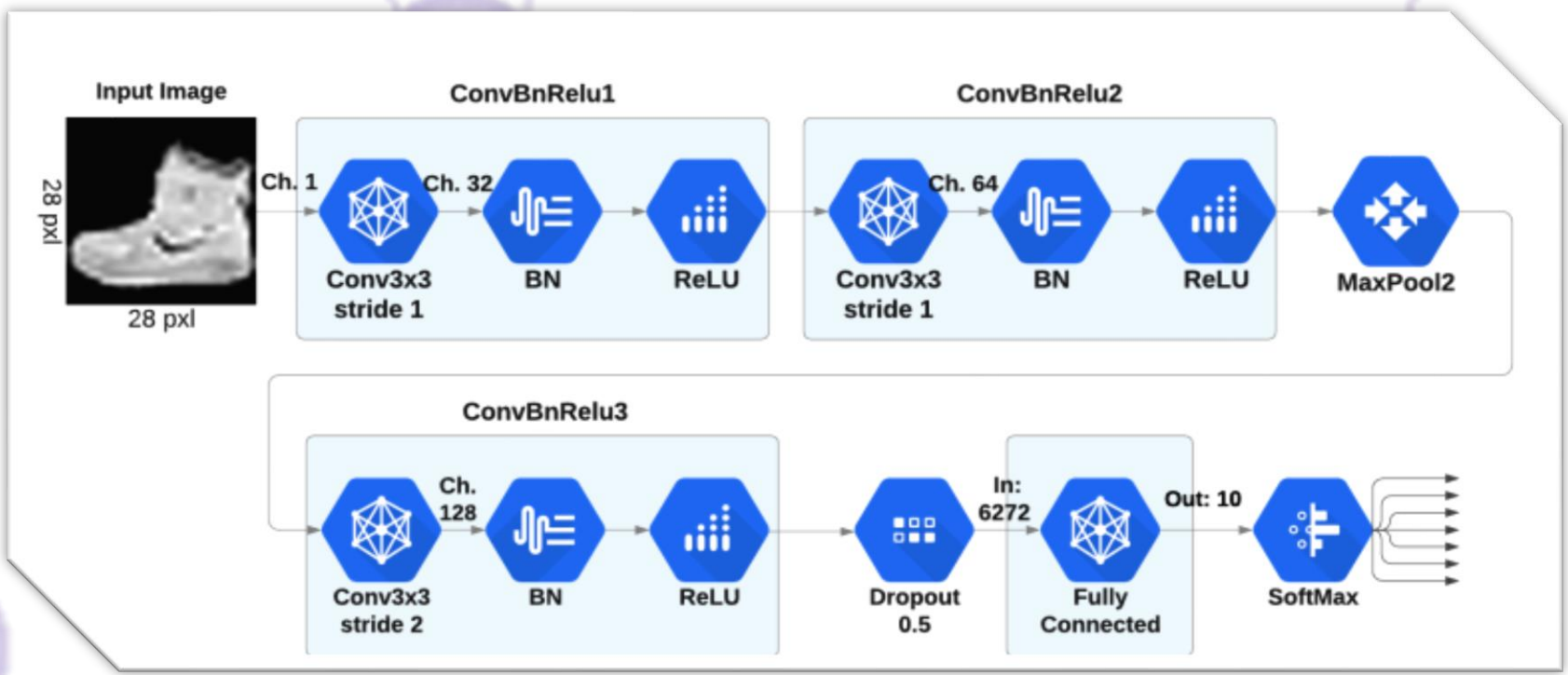


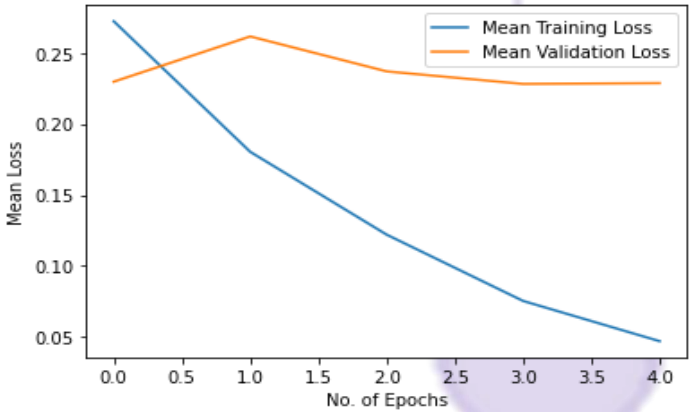
Model	+ Transpose		- Transpose	
	Accuracy	F1	Accuracy	F1
Bert-base + GRU	63%	72%	91%	91%
DistilBert + GRU	51%	65%	78%	76%
Albert + GRU	76%	76%	92%	92%

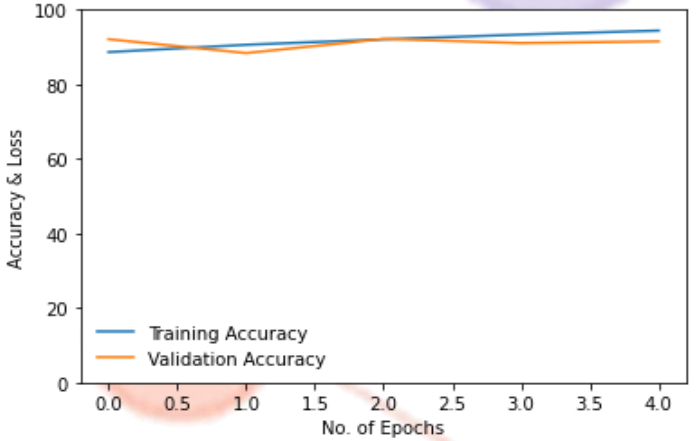
Looking at the confusion matrix I noticed the models were performing badly in classifying entailment of the two arguments when they are labeled as 1 [i.e., “Not next argument”].

I found out that most of the prediction error in the dataset were the transpose of the arguments as label 1. Creating confusion or noise during learning.

5. What could be done next







	precision	recall	f1-score	support
IS NEXT ARGUMENT	0.89	0.93	0.91	5143
IS NOT NEXT ARGUMENT	0.93	0.89	0.91	5142
accuracy			0.91	10285
macro avg	0.91	0.91	0.91	10285
weighted avg	0.91	0.91	0.91	10285

What makes this structure special is that the loss is much lower compared to the other approaches. Adjusting the parameters and adding more noise could help prevent the network to memorize the training data and generalize even better.

Thank You!

The End