240322

* Fixing time series – post length (low p value)
  + ARIMA
  + New algorithm? TS…? - facebook times series neural network
* Ideas on New Measurement for time series – Post Trend, post engagement

> what formula should we use to ensure accuracy and credibility?

> send maksim an email

**Engagement rate by reach:**

**ERR = total num of eng per post / reach per post \* 100**

Post\_trend = (post text length + post title length) / Total text length > make this a ratio

Post\_engagement = total comment + score + (post length \* weight) + upvote ratio > make tis a ratio

Don’t make combine feature > hard to make

Should be using ML that pulls this > feature metrics

If output is yes or no

Regressor: output is a real number > p-value

Multivariate time series data: look at papers and methods

Upvote ratio is already a ratio

* Engagement: positive vs negative posts with how many average likes?
* Question: should we do it separately or with subreddits or all together?

>

> one dataset is a list of dataframes > fix it ? ask chatgpt

* PyTorch vs Tensorflow
  + Tensorflow : working on depression dataset
    - Worked on training BERT both pre and post together failed (shape of pre and post is different)
    - Separately working on pre / post
    - Trial version need to improve the model
* Paper: descriptive analysis, text analysis, statistical testing (time series)
  + Is this good for the paper: yes; 12-15 pages single spaced; formatted like a research paper/conference submission – do local/virtual conference? (computational social conference)
  + Do ppt presentation (15 mins)

240327

Class meeting:

* Tensorflow : just finished up the shape and working on model
* Need to increase our accuracy from building up the model
* Time series????

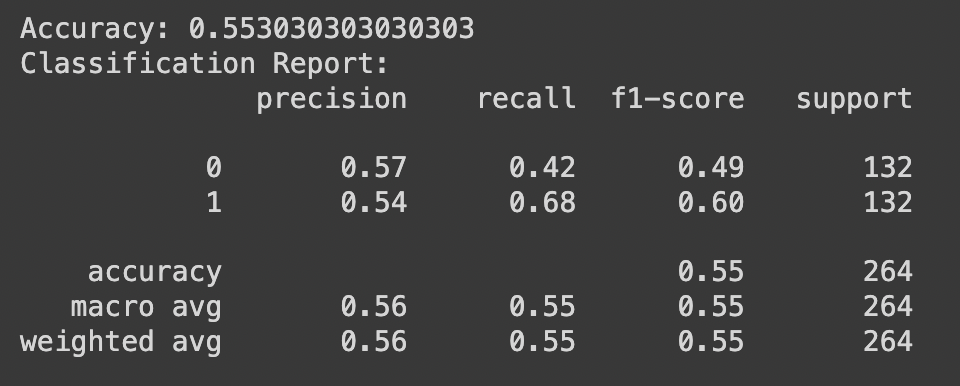
Two different measurements:

Post trend: (post text length + post title length) / Total text length

Post engagement: using total comment, upvote ratio and score → make all measurements as a ratio

Apr 17

* Tensorflow hub and text not working (any suggestions?)
* Pytorch accuracy



Using all dataset is fine?

Wor2vec or bert

If i get higher accuracy should i use tensorflow

Is accuracy relly matte

Distribution - binary cross entropy?

Gpt model - lda

Bert - transformer

T is transformer so same gpt 1

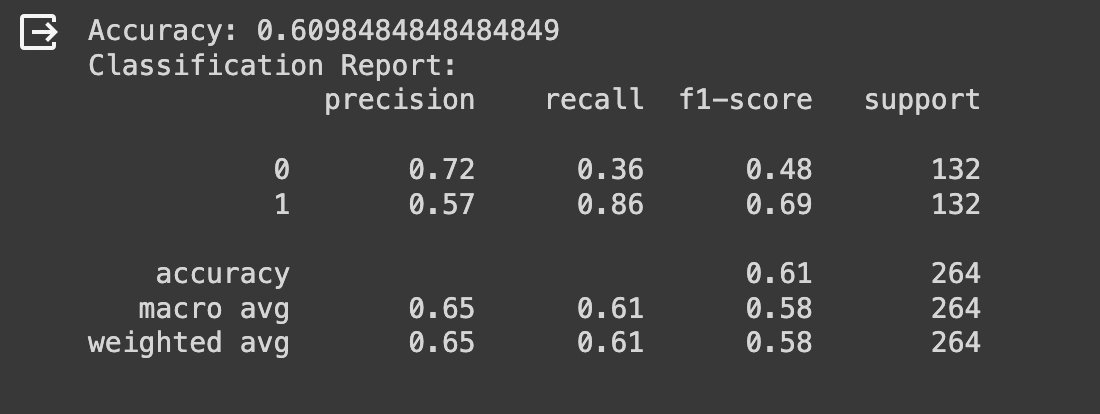
Gpt 3 is 21- hugging face

Bert - Gpt 1 or 3

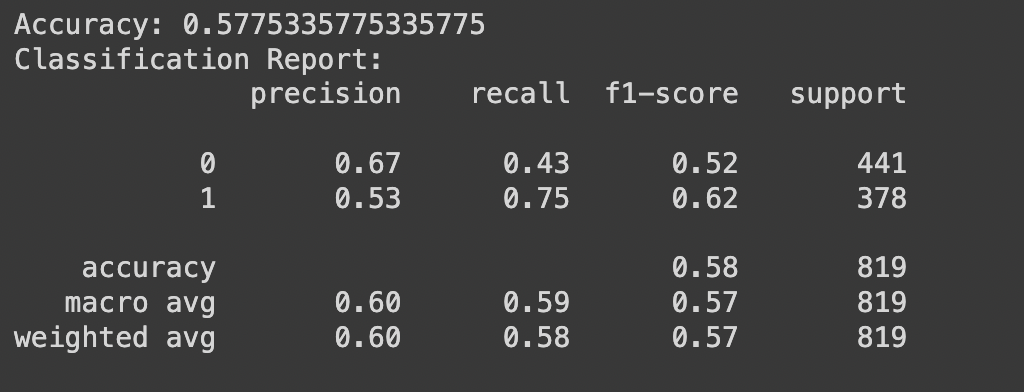
* Dp and anx

Num ep: 3

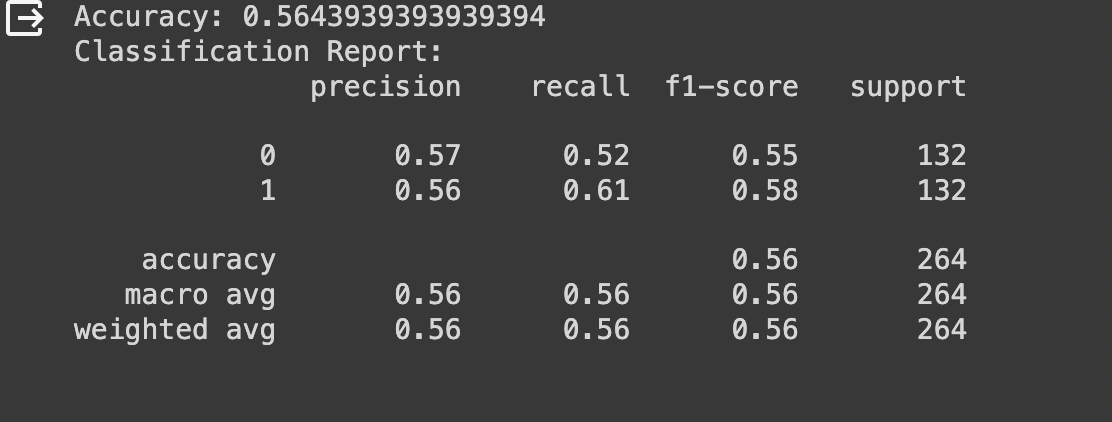
Batch: 16



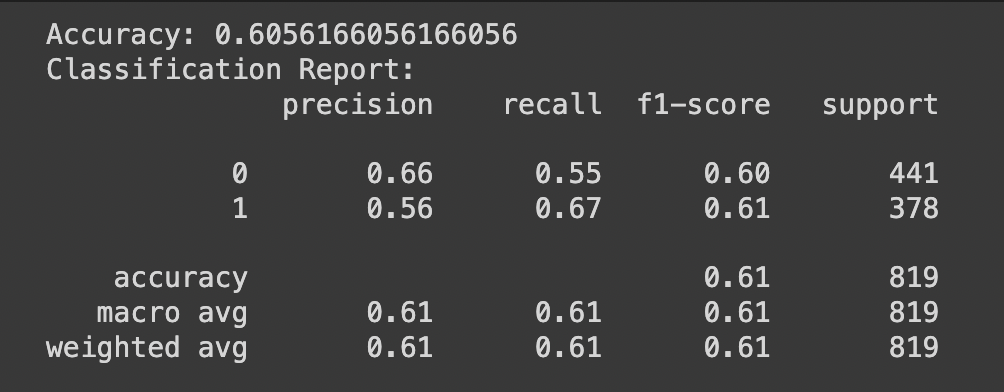
* All Ep:3 Batch:16

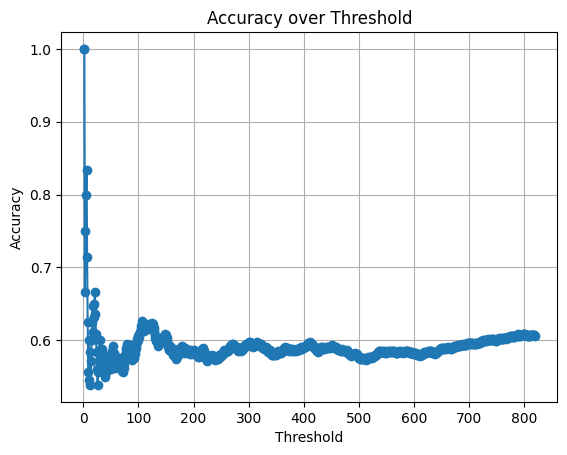


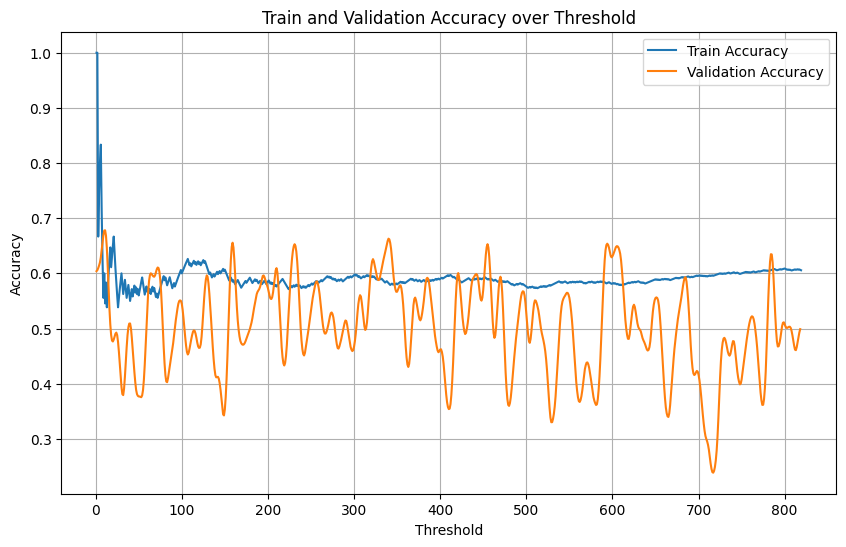
* Ep6, batch 16



* Ep3 ,bat8 all pyt trail final



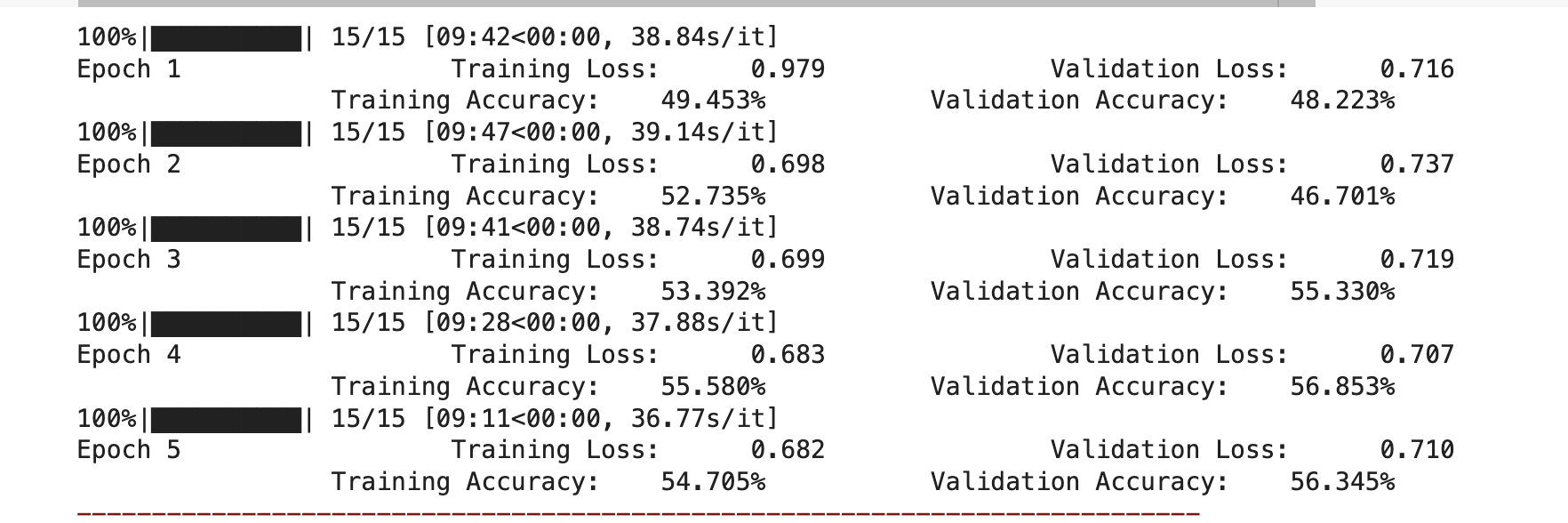


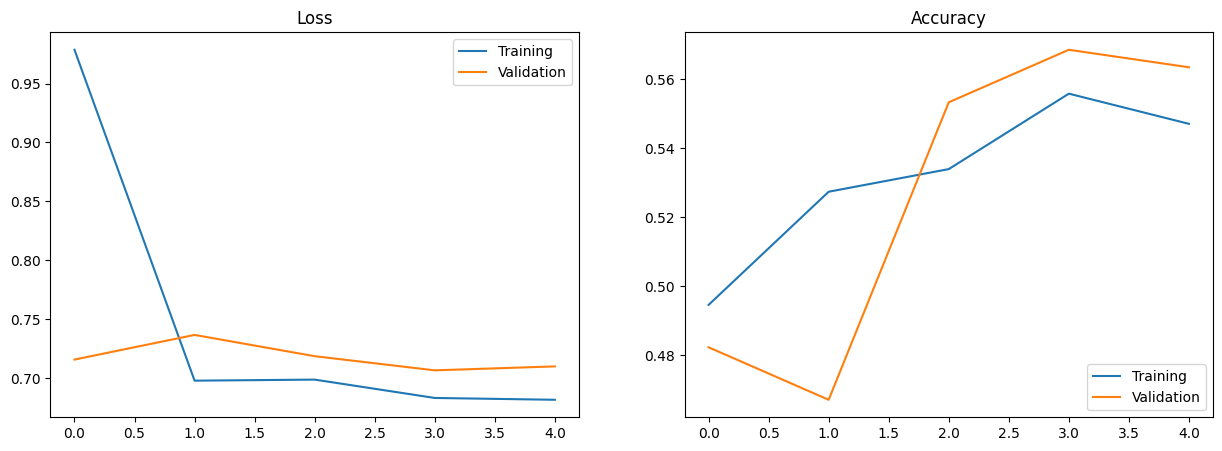


* Ep 5, bat15 medium copy dp anx

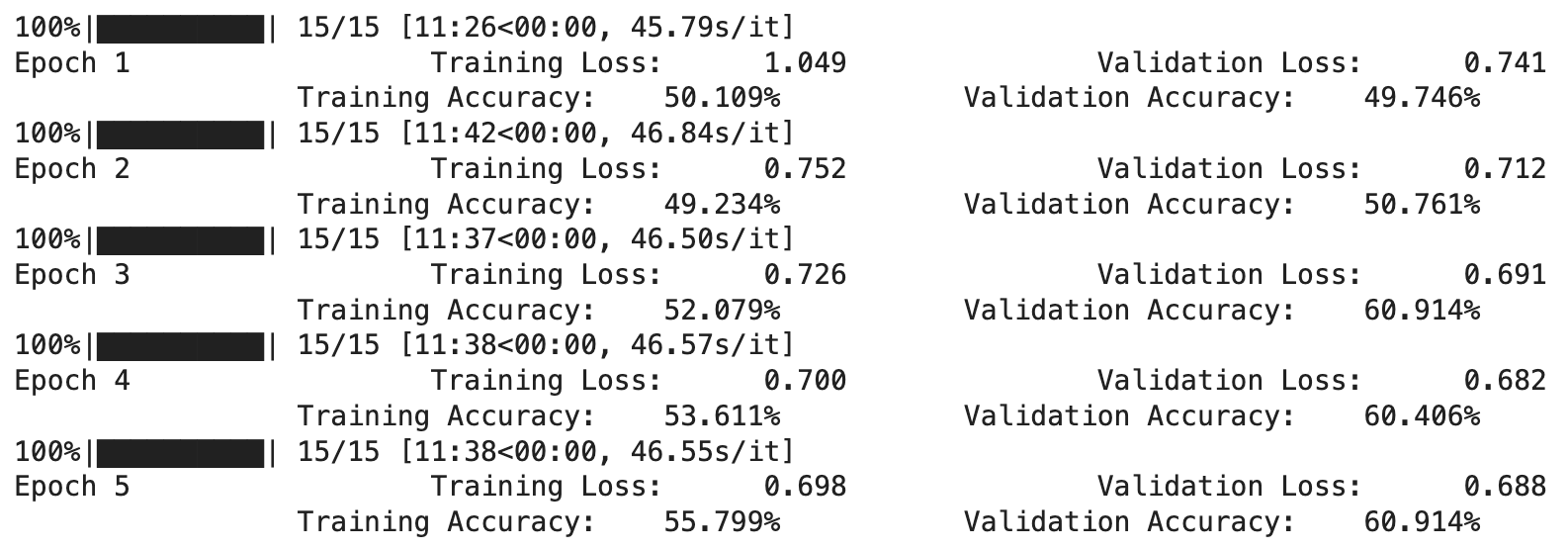
<https://medium.com/@frederik.vl/interpreting-training-validation-accuracy-and-loss-cf16f0d5329f>

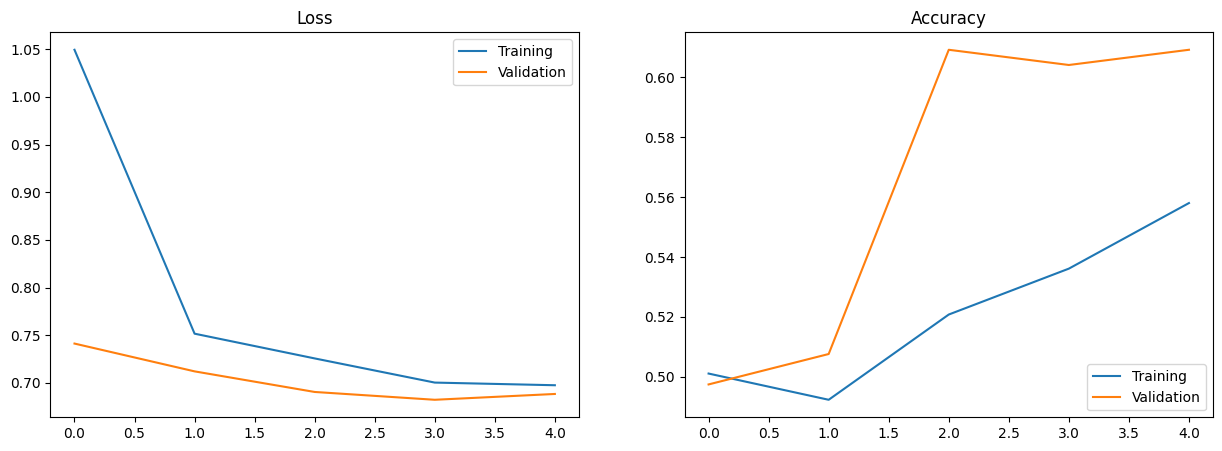
For graph



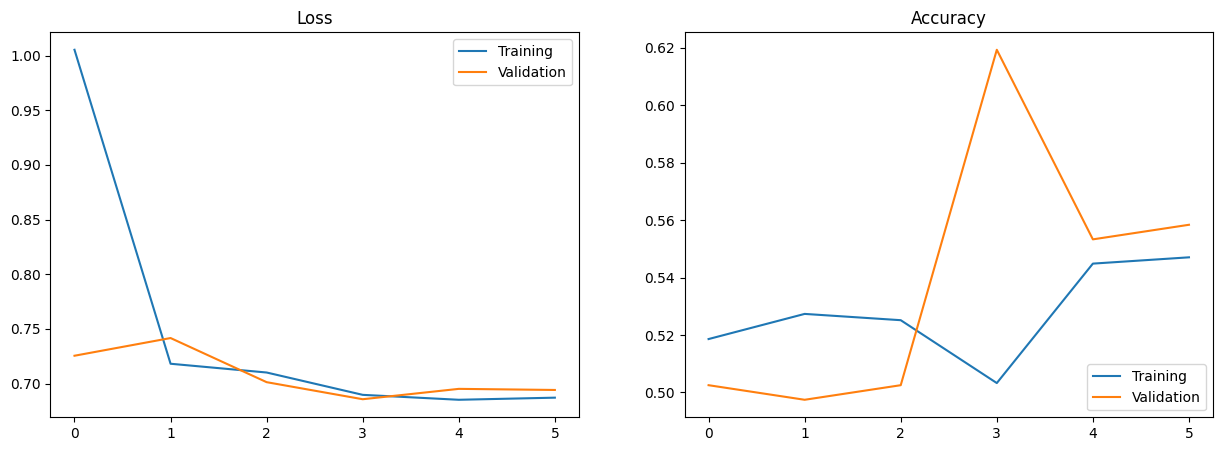
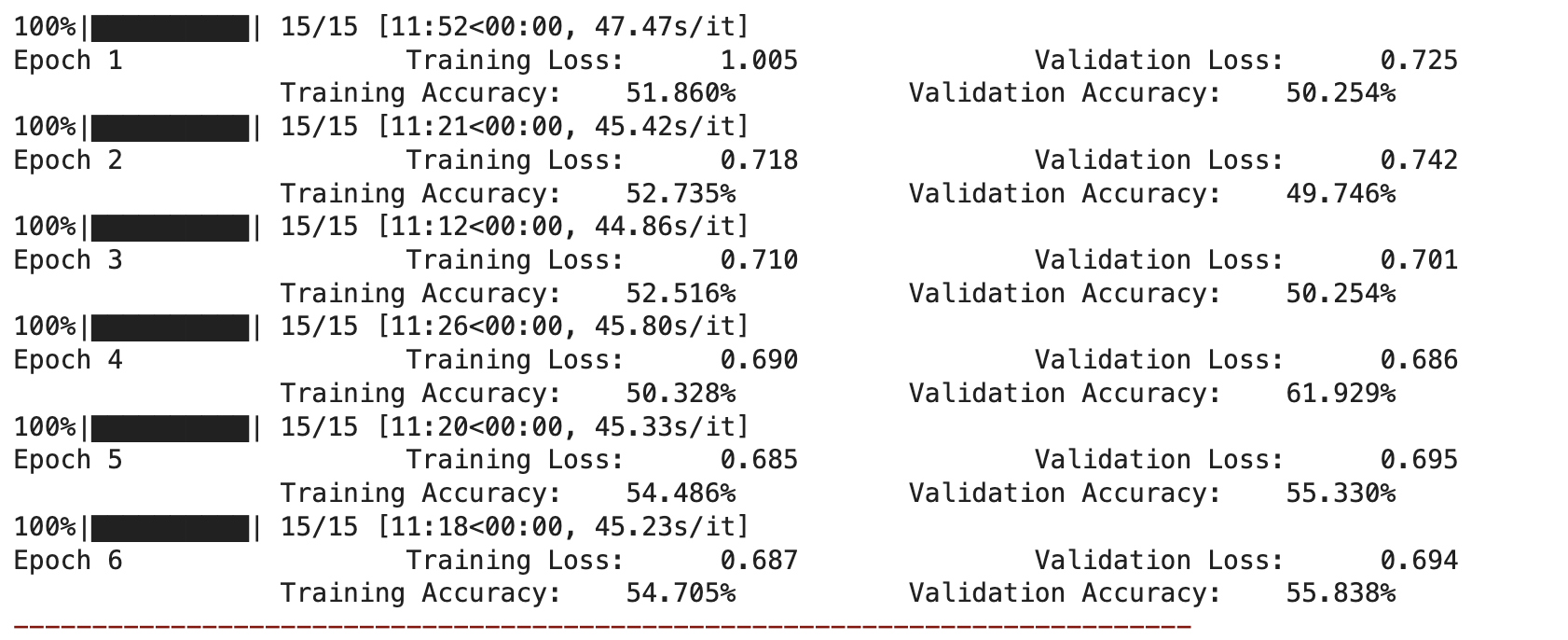


* ep 5, bat 15, dropout2, dp anx

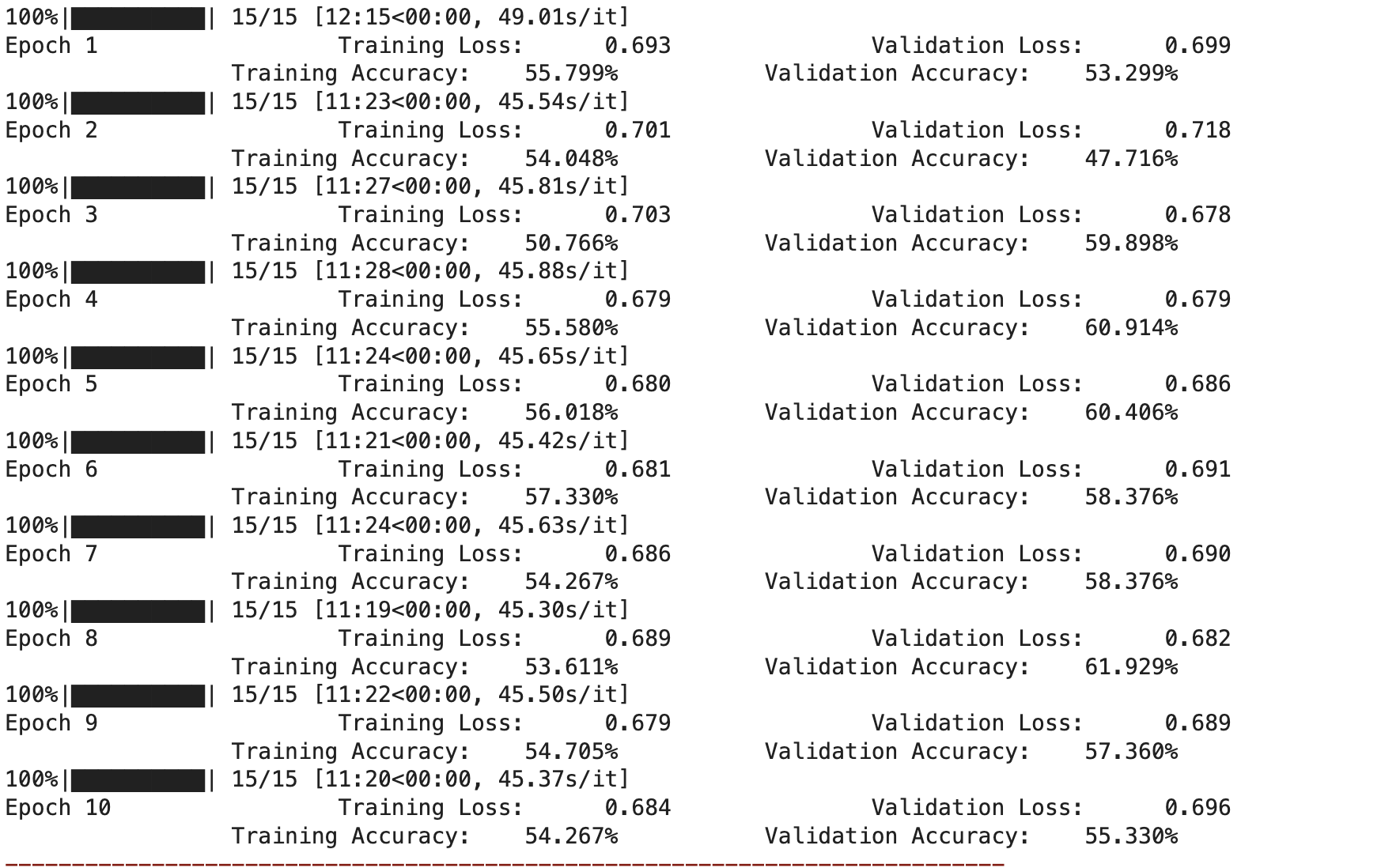


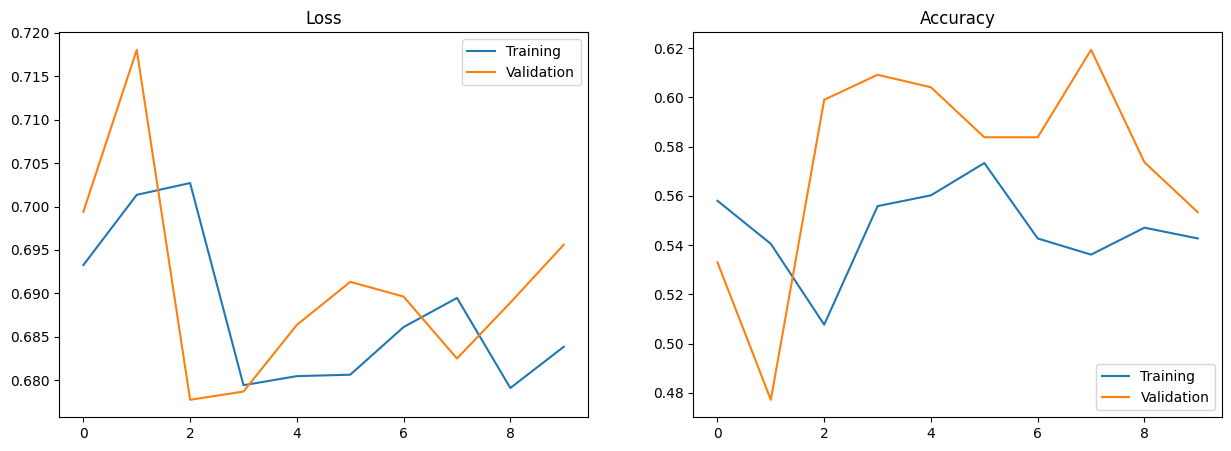


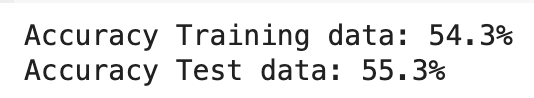
* Ep6, bat15, drop0.3 one dropout, dpanx



-ep10, bat15, dropout0.3, dpanx



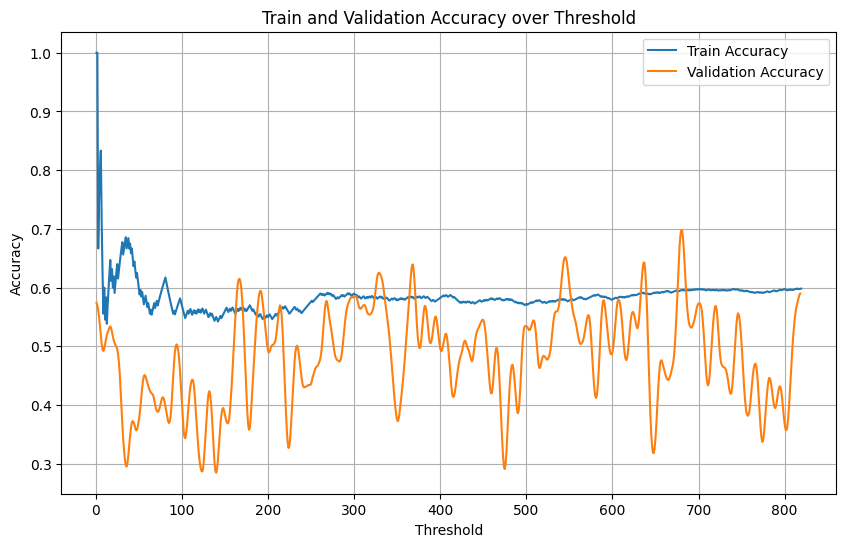


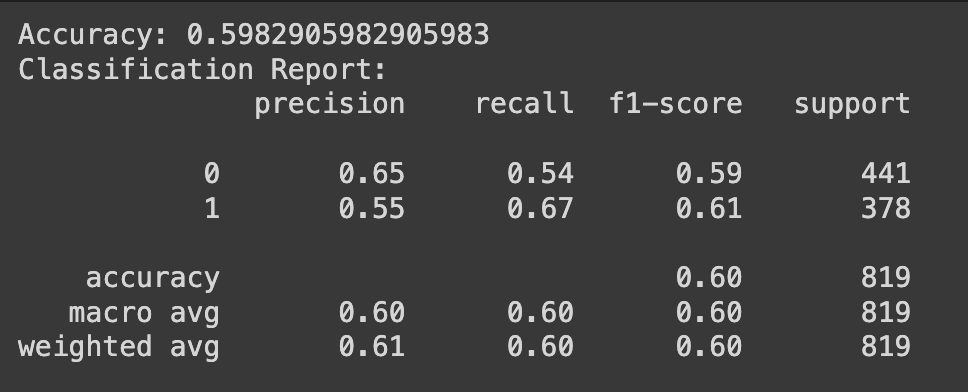


**FINAL MODEL**

**EP5, Bat15 droupout 0.3 all - medium**

**Pyt final 61 all with graph added -ep3 bat15**

****

****

**Pyt final 61 all with graph added -ep3 bat8**

**By Monday**

* Gpt try
* Pyt bat 8, medium
* Paper, medium post

monday

Medium - eda, rewrite

Paper- intro, data, related work, method

Tuesday

Medium done

Paper done

ppt , flyer

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10021013>

<https://ojs.aaai.org/index.php/ICWSM/article/view/14526/14375>

<https://aclanthology.org/W14-3207.pdf>

<https://aclanthology.org/D19-5542.pdf>

usefulsource

<https://www.google.com/search?q=mark+dredze&oq=mark+dredze&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTIHCAEQABiABDIHCAIQABiABDIICAMQABgWGB4yCggEEAAYDxgWGB4yCggFEAAYDxgWGB7SAQkxNTY2NmoxajGoAgCwAgA&sourceid=chrome&ie=UTF-8>

<https://www.google.com/search?q=munmun+de+choudhury+google+scholar&oq=munmun+de+&gs_lcrp=EgZjaHJvbWUqBwgCEAAYgAQyBggAEEUYOTIHCAEQLhiABDIHCAIQABiABDIICAMQABgWGB4yBggEEAUYLNIBCDg1NDZqMGoxqAIAsAIA&sourceid=chrome&ie=UTF-8>

<https://www.google.com/search?q=resnick+maryland&oq=resnick+maryland&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTIICAEQABgWGB4yCAgCEAAYFhgeMggIAxAAGBYYHjIICAQQABgWGB7SAQg5NDY5ajBqMagCALACAA&sourceid=chrome&ie=UTF-8>

<https://perspectiveapi.com/>

<https://github.com/kharrigian/mental-health-datasets>

<https://tweetsets.library.gwu.edu/datasets>