Report 1

I have used IMDB dataset downloaded from hugging face through API. I have applied three models that are pretrained on that dataset. It is a binary Classification Problem.

The models are given as follows:

1. aychang/roberta-base-imdb
2. lvwerra/distilbert-imdb
3. fabriceyhc/bert-base-uncased-imdb
   1. a. Summarize accuracy of all models for their corresponding datasets
4. 1)lvwerra/distilbert-imdb For this getting accuracy of 87%

2) aychang/roberta-base-imdb For this getting average accuracy of 85%

3) fabriceyhc/bert-base-uncased-imdb For this getting average accuracy of 86%

**Results from Roberta**

[{'label': 'pos', 'score': 0.9990511536598206}]

[{'label': 'pos', 'score': 0.9979436993598938}]

[{'label': 'neg', 'score': 0.9984689354896545}]

[{'label': 'pos', 'score': 0.9990767240524292}]

[{'label': 'neg', 'score': 0.9984545707702637}]

[{'label': 'pos', 'score': 0.9991511106491089}]

[{'label': 'pos', 'score': 0.9991419315338135}]

[{'label': 'neg', 'score': 0.9984690546989441}]

[{'label': 'neg', 'score': 0.9984028935432434}]

[{'label': 'pos', 'score': 0.9991521835327148}]

[{'label': 'pos', 'score': 0.9991003274917603}]

[{'label': 'neg', 'score': 0.998471200466156}]

[{'label': 'neg', 'score': 0.9811892509460449}]

[{'label': 'neg', 'score': 0.9690533876419067}]

[{'label': 'pos', 'score': 0.9990980625152588}]

[{'label': 'pos', 'score': 0.9991555213928223}]

[{'label': 'pos', 'score': 0.5810378789901733}]

[{'label': 'pos', 'score': 0.9062329530715942}]

[{'label': 'pos', 'score': 0.99915611743927}]

[{'label': 'pos', 'score': 0.9991461038589478}]

[{'label': 'neg', 'score': 0.9984654188156128}]

[{'label': 'pos', 'score': 0.9686173796653748}]

[{'label': 'neg', 'score': 0.9757689833641052}]

[{'label': 'pos', 'score': 0.9991338849067688}]

[{'label': 'neg', 'score': 0.9933618307113647}]

[{'label': 'pos', 'score': 0.9989452958106995}]

[{'label': 'neg', 'score': 0.998375415802002}]

[{'label': 'pos', 'score': 0.998448371887207}]

[{'label': 'pos', 'score': 0.999147891998291}]

[{'label': 'pos', 'score': 0.9991625547409058}]

[{'label': 'neg', 'score': 0.9983887672424316}]

[{'label': 'neg', 'score': 0.9984686970710754}]

[{'label': 'pos', 'score': 0.9990930557250977}]

[{'label': 'neg', 'score': 0.998430073261261}]

[{'label': 'neg', 'score': 0.9853281378746033}]

[{'label': 'pos', 'score': 0.9991428852081299}]

[{'label': 'neg', 'score': 0.9983469247817993}]

[{'label': 'neg', 'score': 0.9948344230651855}]

[{'label': 'neg', 'score': 0.9847308397293091}]

[{'label': 'pos', 'score': 0.9650424122810364}]

[{'label': 'neg', 'score': 0.9984708428382874}]

[{'label': 'neg', 'score': 0.9984715580940247}]

[{'label': 'neg', 'score': 0.9982732534408569}]

[{'label': 'neg', 'score': 0.9984702467918396}]

[{'label': 'neg', 'score': 0.9984556436538696}]

[{'label': 'pos', 'score': 0.9775913953781128}]

[{'label': 'neg', 'score': 0.9912468791007996}]

[{'label': 'pos', 'score': 0.9518259763717651}]

[{'label': 'pos', 'score': 0.9991315007209778}]

[{'label': 'neg', 'score': 0.9908527731895447}]

[{'label': 'pos', 'score': 0.999076247215271}]

[{'label': 'pos', 'score': 0.9991497993469238}]

[{'label': 'neg', 'score': 0.9984606504440308}]

[{'label': 'neg', 'score': 0.9956193566322327}]

[{'label': 'pos', 'score': 0.9991031885147095}]

[{'label': 'pos', 'score': 0.9946111440658569}]

[{'label': 'neg', 'score': 0.9984704852104187}]

[{'label': 'pos', 'score': 0.9661308526992798}]

[{'label': 'pos', 'score': 0.9982230067253113}]

[{'label': 'neg', 'score': 0.9827708005905151}]

[{'label': 'neg', 'score': 0.9977211356163025}]

[{'label': 'neg', 'score': 0.9980918765068054}]

[{'label': 'pos', 'score': 0.9988560676574707}]

[{'label': 'neg', 'score': 0.9984007477760315}]

[{'label': 'pos', 'score': 0.9991613626480103}]

[{'label': 'neg', 'score': 0.9984710812568665}]

[{'label': 'neg', 'score': 0.9978773593902588}]

[{'label': 'pos', 'score': 0.9991300702095032}]

[{'label': 'pos', 'score': 0.9989994168281555}]

[{'label': 'neg', 'score': 0.99847012758255}]

[{'label': 'pos', 'score': 0.9987515211105347}]

[{'label': 'neg', 'score': 0.9984013438224792}]

[{'label': 'neg', 'score': 0.998420238494873}]

[{'label': 'pos', 'score': 0.9991484880447388}]

[{'label': 'pos', 'score': 0.9931150078773499}]

[{'label': 'neg', 'score': 0.9849397540092468}]

[{'label': 'neg', 'score': 0.9984684586524963}]

[{'label': 'pos', 'score': 0.995184600353241}]

[{'label': 'neg', 'score': 0.9980652928352356}]

[{'label': 'neg', 'score': 0.9915174245834351}]

[{'label': 'neg', 'score': 0.9984094500541687}]

[{'label': 'pos', 'score': 0.998877227306366}]

[{'label': 'neg', 'score': 0.9950435161590576}]

[{'label': 'pos', 'score': 0.9988172054290771}]

[{'label': 'neg', 'score': 0.9839058518409729}]

[{'label': 'pos', 'score': 0.9991453886032104}]

[{'label': 'pos', 'score': 0.9991101622581482}]

[{'label': 'neg', 'score': 0.9984710812568665}]

[{'label': 'neg', 'score': 0.9903221726417542}]

[{'label': 'pos', 'score': 0.9991551637649536}]

[{'label': 'pos', 'score': 0.9894599914550781}]

[{'label': 'pos', 'score': 0.9978278279304504}]

[{'label': 'neg', 'score': 0.9941161870956421}]

[{'label': 'pos', 'score': 0.9991453886032104}]

[{'label': 'pos', 'score': 0.9606682658195496}]

[{'label': 'neg', 'score': 0.9981095790863037}]

[{'label': 'neg', 'score': 0.9984670281410217}]

[{'label': 'neg', 'score': 0.9981133937835693}]

[{'label': 'pos', 'score': 0.9480404853820801}]

[{'label': 'pos', 'score': 0.9991501569747925}]

0.85

**Results From Bert Uncase**

**[**{'label': 'pos', 'score': 0.9996671676635742}]

[{'label': 'pos', 'score': 0.9996638298034668}]

[{'label': 'neg', 'score': 0.9924436211585999}]

[{'label': 'pos', 'score': 0.9996439218521118}]

[{'label': 'neg', 'score': 0.9996143579483032}]

[{'label': 'pos', 'score': 0.9996660947799683}]

[{'label': 'pos', 'score': 0.9996684789657593}]

[{'label': 'neg', 'score': 0.9996353387832642}]

[{'label': 'neg', 'score': 0.9995144605636597}]

[{'label': 'pos', 'score': 0.9996663331985474}]

[{'label': 'pos', 'score': 0.9996657371520996}]

[{'label': 'neg', 'score': 0.9996557235717773}]

[{'label': 'pos', 'score': 0.9924954175949097}]

[{'label': 'neg', 'score': 0.9961469173431396}]

[{'label': 'pos', 'score': 0.9996672868728638}]

[{'label': 'pos', 'score': 0.9996678829193115}]

[{'label': 'pos', 'score': 0.9955743551254272}]

[{'label': 'pos', 'score': 0.9931942224502563}]

[{'label': 'pos', 'score': 0.9996693134307861}]

[{'label': 'pos', 'score': 0.9996651411056519}]

[{'label': 'neg', 'score': 0.9995872378349304}]

[{'label': 'pos', 'score': 0.9996141195297241}]

[{'label': 'pos', 'score': 0.985292375087738}]

[{'label': 'pos', 'score': 0.9996674060821533}]

[{'label': 'neg', 'score': 0.9995747208595276}]

[{'label': 'pos', 'score': 0.9996600151062012}]

[{'label': 'neg', 'score': 0.999234676361084}]

[{'label': 'pos', 'score': 0.9993370175361633}]

[{'label': 'pos', 'score': 0.9996693134307861}]

[{'label': 'pos', 'score': 0.9996629953384399}]

[{'label': 'neg', 'score': 0.999548614025116}]

[{'label': 'neg', 'score': 0.9995296001434326}]

[{'label': 'pos', 'score': 0.9995926022529602}]

[{'label': 'neg', 'score': 0.9991829991340637}]

[{'label': 'neg', 'score': 0.9989312291145325}]

[{'label': 'pos', 'score': 0.9996080994606018}]

[{'label': 'neg', 'score': 0.9961207509040833}]

[{'label': 'pos', 'score': 0.9971452355384827}]

[{'label': 'neg', 'score': 0.9994895458221436}]

[{'label': 'pos', 'score': 0.9959198236465454}]

[{'label': 'neg', 'score': 0.9987066984176636}]

[{'label': 'neg', 'score': 0.999549925327301}]

[{'label': 'neg', 'score': 0.9996354579925537}]

[{'label': 'neg', 'score': 0.9993453621864319}]

[{'label': 'neg', 'score': 0.9994401335716248}]

[{'label': 'pos', 'score': 0.9996427297592163}]

[{'label': 'neg', 'score': 0.997270405292511}]

[{'label': 'pos', 'score': 0.9978220462799072}]

[{'label': 'pos', 'score': 0.9996669292449951}]

[{'label': 'neg', 'score': 0.9952364563941956}]

[{'label': 'pos', 'score': 0.9996618032455444}]

[{'label': 'pos', 'score': 0.9995837807655334}]

[{'label': 'neg', 'score': 0.9995850920677185}]

[{'label': 'pos', 'score': 0.967881977558136}]

[{'label': 'pos', 'score': 0.9996680021286011}]

[{'label': 'pos', 'score': 0.9981333613395691}]

[{'label': 'neg', 'score': 0.9995703101158142}]

[{'label': 'pos', 'score': 0.9682725667953491}]

[{'label': 'pos', 'score': 0.9986288547515869}]

[{'label': 'neg', 'score': 0.7687053680419922}]

[{'label': 'neg', 'score': 0.999519944190979}]

[{'label': 'neg', 'score': 0.9995952248573303}]

[{'label': 'neg', 'score': 0.9989711046218872}]

[{'label': 'neg', 'score': 0.9994860887527466}]

[{'label': 'pos', 'score': 0.9996668100357056}]

[{'label': 'neg', 'score': 0.9996367692947388}]

[{'label': 'neg', 'score': 0.9994472861289978}]

[{'label': 'pos', 'score': 0.9996665716171265}]

[{'label': 'pos', 'score': 0.9996687173843384}]

[{'label': 'neg', 'score': 0.9996461868286133}]

[{'label': 'pos', 'score': 0.9995123147964478}]

[{'label': 'neg', 'score': 0.9995288848876953}]

[{'label': 'neg', 'score': 0.9996185302734375}]

[{'label': 'pos', 'score': 0.9996479749679565}]

[{'label': 'neg', 'score': 0.950439989566803}]

[{'label': 'pos', 'score': 0.5945958495140076}]

[{'label': 'neg', 'score': 0.9995146989822388}]

[{'label': 'pos', 'score': 0.9996672868728638}]

[{'label': 'neg', 'score': 0.999380350112915}]

[{'label': 'pos', 'score': 0.9994620680809021}]

[{'label': 'neg', 'score': 0.9995951056480408}]

[{'label': 'pos', 'score': 0.9996212720870972}]

[{'label': 'neg', 'score': 0.9978680610656738}]

[{'label': 'pos', 'score': 0.9996597766876221}]

[{'label': 'pos', 'score': 0.9993077516555786}]

[{'label': 'pos', 'score': 0.999667763710022}]

[{'label': 'pos', 'score': 0.999657392501831}]

[{'label': 'neg', 'score': 0.9996325969696045}]

[{'label': 'neg', 'score': 0.9989340901374817}]

[{'label': 'pos', 'score': 0.9996672868728638}]

[{'label': 'neg', 'score': 0.960587739944458}]

[{'label': 'pos', 'score': 0.9996445178985596}]

[{'label': 'neg', 'score': 0.9994874000549316}]

[{'label': 'pos', 'score': 0.9996562004089355}]

[{'label': 'pos', 'score': 0.9992489218711853}]

[{'label': 'pos', 'score': 0.9984129667282104}]

[{'label': 'neg', 'score': 0.9994899034500122}]

[{'label': 'neg', 'score': 0.9995846152305603}]

[{'label': 'pos', 'score': 0.999382734298706}]

[{'label': 'pos', 'score': 0.999664306640625}]

0.86

**Results From DistilBert**:

[{'label': 'POSITIVE', 'score': 0.959995448589325}]

[{'label': 'POSITIVE', 'score': 0.7070561051368713}]

[{'label': 'NEGATIVE', 'score': 0.9883654117584229}]

[{'label': 'POSITIVE', 'score': 0.9629437923431396}]

[{'label': 'NEGATIVE', 'score': 0.9923607707023621}]

[{'label': 'POSITIVE', 'score': 0.9937682151794434}]

[{'label': 'POSITIVE', 'score': 0.9853847622871399}]

[{'label': 'NEGATIVE', 'score': 0.9958129525184631}]

[{'label': 'NEGATIVE', 'score': 0.9756316542625427}]

[{'label': 'POSITIVE', 'score': 0.9933367967605591}]

[{'label': 'POSITIVE', 'score': 0.9929167032241821}]

[{'label': 'NEGATIVE', 'score': 0.9961766004562378}]

[{'label': 'NEGATIVE', 'score': 0.6506089568138123}]

[{'label': 'NEGATIVE', 'score': 0.9035837054252625}]

[{'label': 'POSITIVE', 'score': 0.9911532998085022}]

[{'label': 'POSITIVE', 'score': 0.994312047958374}]

[{'label': 'POSITIVE', 'score': 0.5728194713592529}]

[{'label': 'NEGATIVE', 'score': 0.7526023983955383}]

[{'label': 'POSITIVE', 'score': 0.9936975836753845}]

[{'label': 'POSITIVE', 'score': 0.9582812190055847}]

[{'label': 'NEGATIVE', 'score': 0.9944151639938354}]

[{'label': 'POSITIVE', 'score': 0.9677115082740784}]

[{'label': 'NEGATIVE', 'score': 0.5504311919212341}]

[{'label': 'POSITIVE', 'score': 0.9877728819847107}]

[{'label': 'NEGATIVE', 'score': 0.9914381504058838}]

[{'label': 'POSITIVE', 'score': 0.9511362314224243}]

[{'label': 'NEGATIVE', 'score': 0.9721881151199341}]

[{'label': 'POSITIVE', 'score': 0.8292925953865051}]

[{'label': 'POSITIVE', 'score': 0.9909296631813049}]

[{'label': 'POSITIVE', 'score': 0.9957807064056396}]

[{'label': 'NEGATIVE', 'score': 0.9898263216018677}]

[{'label': 'NEGATIVE', 'score': 0.9574320316314697}]

[{'label': 'POSITIVE', 'score': 0.5946581959724426}]

[{'label': 'NEGATIVE', 'score': 0.9916770458221436}]

[{'label': 'NEGATIVE', 'score': 0.6160079836845398}]

[{'label': 'POSITIVE', 'score': 0.9868042469024658}]

[{'label': 'NEGATIVE', 'score': 0.9844781160354614}]

[{'label': 'NEGATIVE', 'score': 0.9376795887947083}]

[{'label': 'NEGATIVE', 'score': 0.9566333889961243}]

[{'label': 'POSITIVE', 'score': 0.9207740426063538}]

[{'label': 'NEGATIVE', 'score': 0.9954532384872437}]

[{'label': 'NEGATIVE', 'score': 0.99595046043396}]

[{'label': 'NEGATIVE', 'score': 0.9956703186035156}]

[{'label': 'NEGATIVE', 'score': 0.9922086000442505}]

[{'label': 'NEGATIVE', 'score': 0.9906340837478638}]

[{'label': 'POSITIVE', 'score': 0.9128568172454834}]

[{'label': 'NEGATIVE', 'score': 0.5274920463562012}]

[{'label': 'NEGATIVE', 'score': 0.8232696652412415}]

[{'label': 'POSITIVE', 'score': 0.9888685941696167}]

[{'label': 'NEGATIVE', 'score': 0.9744884371757507}]

[{'label': 'POSITIVE', 'score': 0.9581727385520935}]

[{'label': 'POSITIVE', 'score': 0.96431964635849}]

[{'label': 'NEGATIVE', 'score': 0.978840708732605}]

[{'label': 'NEGATIVE', 'score': 0.6679476499557495}]

[{'label': 'POSITIVE', 'score': 0.9885507225990295}]

[{'label': 'POSITIVE', 'score': 0.9429939389228821}]

[{'label': 'NEGATIVE', 'score': 0.9950234889984131}]

[{'label': 'POSITIVE', 'score': 0.8822948336601257}]

[{'label': 'NEGATIVE', 'score': 0.8815016746520996}]

[{'label': 'NEGATIVE', 'score': 0.9404389262199402}]

[{'label': 'NEGATIVE', 'score': 0.9880489706993103}]

[{'label': 'NEGATIVE', 'score': 0.907969057559967}]

[{'label': 'POSITIVE', 'score': 0.9318650364875793}]

[{'label': 'NEGATIVE', 'score': 0.9598066210746765}]

[{'label': 'POSITIVE', 'score': 0.9926040768623352}]

[{'label': 'NEGATIVE', 'score': 0.9943320751190186}]

[{'label': 'NEGATIVE', 'score': 0.9545159935951233}]

[{'label': 'POSITIVE', 'score': 0.992954671382904}]

[{'label': 'POSITIVE', 'score': 0.9855681657791138}]

[{'label': 'NEGATIVE', 'score': 0.9932271838188171}]

[{'label': 'POSITIVE', 'score': 0.9842946529388428}]

[{'label': 'NEGATIVE', 'score': 0.9894519448280334}]

[{'label': 'NEGATIVE', 'score': 0.9966416358947754}]

[{'label': 'POSITIVE', 'score': 0.9892275333404541}]

[{'label': 'POSITIVE', 'score': 0.7823293805122375}]

[{'label': 'NEGATIVE', 'score': 0.7725049257278442}]

[{'label': 'NEGATIVE', 'score': 0.9958641529083252}]

[{'label': 'POSITIVE', 'score': 0.9621001482009888}]

[{'label': 'NEGATIVE', 'score': 0.7387871146202087}]

[{'label': 'POSITIVE', 'score': 0.9458786249160767}]

[{'label': 'NEGATIVE', 'score': 0.9913930892944336}]

[{'label': 'POSITIVE', 'score': 0.9861375689506531}]

[{'label': 'NEGATIVE', 'score': 0.8657177686691284}]

[{'label': 'POSITIVE', 'score': 0.8966444134712219}]

[{'label': 'POSITIVE', 'score': 0.7515055537223816}]

[{'label': 'POSITIVE', 'score': 0.9934735894203186}]

[{'label': 'POSITIVE', 'score': 0.9911075234413147}]

[{'label': 'NEGATIVE', 'score': 0.9951316118240356}]

[{'label': 'NEGATIVE', 'score': 0.8707504868507385}]

[{'label': 'POSITIVE', 'score': 0.9941307902336121}]

[{'label': 'NEGATIVE', 'score': 0.9102921485900879}]

[{'label': 'POSITIVE', 'score': 0.9820427298545837}]

[{'label': 'NEGATIVE', 'score': 0.9938855767250061}]

[{'label': 'POSITIVE', 'score': 0.9264296889305115}]

[{'label': 'POSITIVE', 'score': 0.6156538724899292}]

[{'label': 'NEGATIVE', 'score': 0.945882260799408}]

[{'label': 'NEGATIVE', 'score': 0.9920491576194763}]

[{'label': 'NEGATIVE', 'score': 0.9831525087356567}]

[{'label': 'POSITIVE', 'score': 0.8347100615501404}]

[{'label': 'POSITIVE', 'score': 0.9954946041107178}]

0.87

1. 1. b. Write out explicit comparisons and discussion of results (were there any surprises? Any possible bugs that could be causing problems?)

All the pertained models are doing equally good. All the accuracies are near about.

The bug I have found while applying is that bert model have tokenization length of 512 if the text is longer than that after tokenization it shows error because if the word is not in vocab of the bert it will try to tokenize it chunk wise making it’s the length of text after tokenization more.

* 1. c. Small section on limitations of your comparisons

limitations is that testing result is on small dataset as compared to the proportion of training data. No hyperparameter tuning has been performed which could improve the performance. As models are pretrained, no training has been performed on train data.

* 1. d. Small section on how you could use these pretrained models in the future if you were to do research in these directions

Pretrained models are very useful as these save a lot of time in retraining. I will seek that my research dataset is resembling with nature of dataset that is available in hugging face, I will seek pretrained models on that dataset, and I will use that on my research dataset.