**Laporan Hasil Eksperimen Machine Learning untuk Cyberbullying Detection**

Kelompok 6 :

Albert Yang - 2602135912

Griventh Griffith Agustin - 2602103021

Joshua Kenneth Van Dyon - 2602136272

1. Berikut Tabel Komparasi hasil dari 3 machine learning yang kami teliti :

| Random Forest | | | | |
| --- | --- | --- | --- | --- |
|  | precision | Recall | F1- Score | Support |
| age | 0.93294 | 0.98311 | 0.95737 | 1599 |
| ethnicity | 0.96496 | 0.96556 | 0.96526 | 1597 |
| gender | 0.89509 | 0.78804 | 0.83816 | 1505 |
| not\_cyberbullying | 0.56957 | 0.49078 | 0.52725 | 1518 |
| other\_cyberbullying | 0.55574 | 0.68780 | 0.61476 | 1435 |
| religion | 0.93713 | 0.91438 | 0.92561 | 1565 |
|  |  |  |  |  |
| Accuracy |  |  | 0.80852 | 9219 |
| ROC - AUC | 0.95108 |  |  |  |

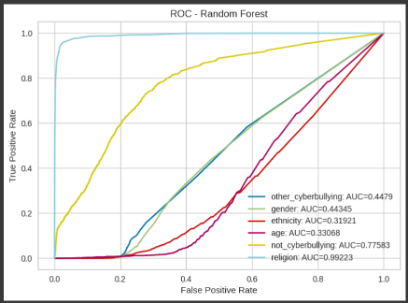
| Recurrent Neural Network | | | | |
| --- | --- | --- | --- | --- |
|  | precision | Recall | F1- Score | Support |
| age | 0.94750 | 0.97061 | 0.95891 | 1599 |
| ethnicity | 0.96728 | 0.96243 | 0.96485 | 1597 |
| gender | 0.86058 | 0.83256 | 0.84634 | 1505 |
| not\_cyberbullying | 0.55307 | 0.43939 | 0.48972 | 1518 |
| other\_cyberbullying | 0.52972 | 0.62718 | 0.57435 | 1518 |
| religion | 0.91416 | 0.95272 | 0.93304 | 1565 |
|  |  |  |  |  |
| Accuracy |  |  | 0.80269 | 9219 |
| ROC - AUC | 0.95654 |  |  |  |

| Support Vector Machine | | | | |
| --- | --- | --- | --- | --- |
|  | precision | Recall | F1- Score | Support |
| age | 0.93468 | 0.98437 | 0.95888 | 1599 |
| ethnicity | 0.96375 | 0.96556 | 0.96465 | 1597 |
| gender | 0.90023 | 0.78538 | 0.83889 | 1505 |
| not\_cyberbullying | 0.57100 | 0.49802 | 0.53202 | 1518 |
| other\_cyberbullying | 0.55824 | 0.69129 | 0.61768 | 1435 |
| religion | 0.93951 | 0.91310 | 0.92612 | 1565 |
|  |  |  |  |  |
| Accuracy |  |  | 0.81083 | 9219 |
| ROC - AUC | 0.96021 |  |  |  |

Setelah diteliti kita mendapatkan akurasi yang cukup sama terhadap 3 machine learning yang saya gunakan dan bisa dilihat dari tabelnya yang paling tinggi adalah support vector machine dengan accuracy 0.81083. Dan ROC AUC yang paling tinggi dari ketiga machine learning yang saya gunakan juga Support Vector Machine.

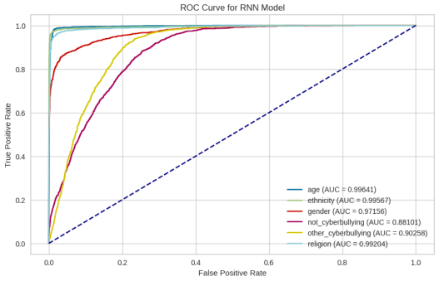
1. ROC Curve :

ROC - Random Forest



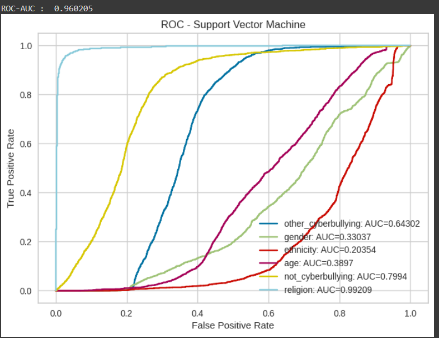
Berdasarkan grafik ROC AUC diatas, algoritma random forest hanya baik dalam memprediksi kelas “religion” dan “not\_cyberbullying” saja.

ROC - Recurrent Neural Network



Berdasarkan grafik ROC AUC diatas, algoritma RNN sangat baik, bahkan diatas garis diagonal dalam memprediksi semua kelas pada dataset.

ROC - Support Vector Machine

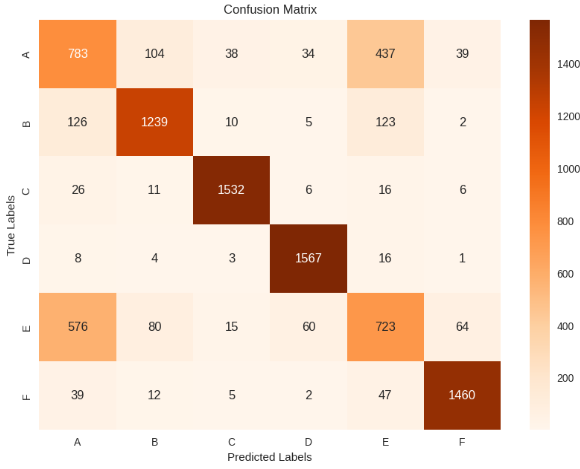


Berdasarkan grafik ROC AUC diatas, algoritma cukup baik dalam memprediksi kelas “religion”, “not\_cyberbullying”, dan “gender” saja, sisanya di bawah garis diagonal.

Berdasarkan ketiga grafik ROC AUC diatas, dapat disimpulkan bahwa model machine learning yang terbaik untuk mengklasifikasikan teks cyberbullying adalah algoritma Recurrent Neural Network atau RNN.

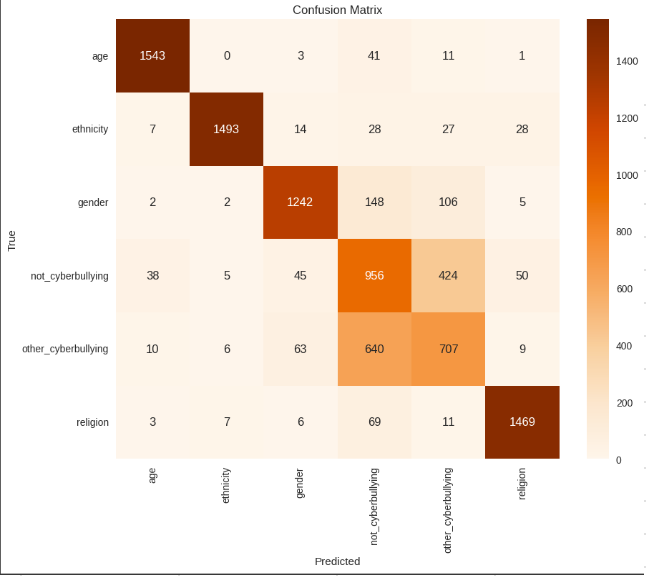
1. Confusion Matrix

Confusion Matrix - Random Forest



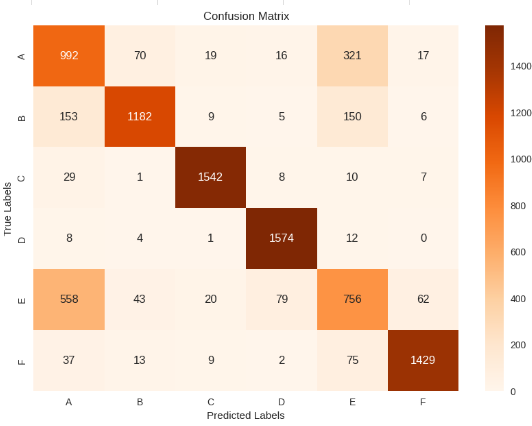
Berdasarkan confusion matrix diatas, dapat dilihat bahwa model Random Forest cukup baik dalam memprediksi kelas B,C,D dan F dengan rentang nilai akurasi 1200 - 1500. Namun dapat dilihat model ini memiliki False Negative di kelas E, serta False positive di kelas E, sehingga model ini dapat kurang akurat.

Confusion Matrix - Recurrent Neural Network



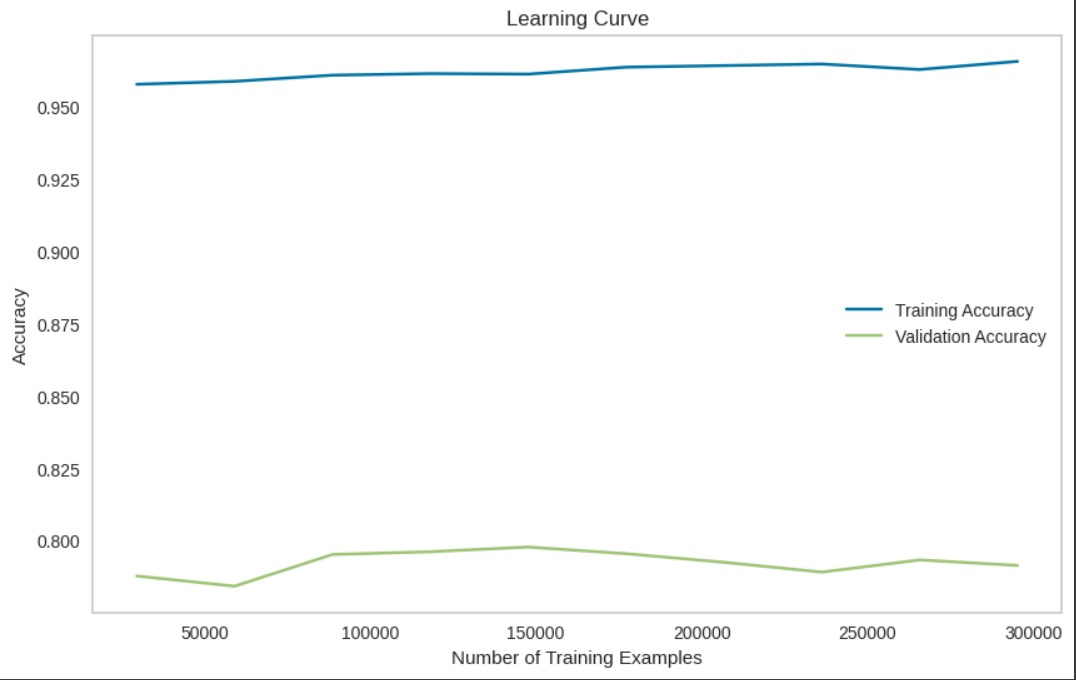
Berdasarkan confusion matrix diatas, dapat dilihat bahwa model RNN sangat baik dalam memprediksi kelas age, ethnicity, gender, not\_cyberbullying, other\_cyberbullying dan religion dengan rentang nilai akurasi 700 - 1500 serta minim False Negative dan True Negative yang signifikan di kelas other\_cyberbullying, ini dapat menjadi indikasi akurasi yang baik.

Confusion Matrix - Support Vector Machine



Berdasarkan confusion matrix diatas, dapat dilihat bahwa model SVM cukup baik dalam memprediksi kelas B,C,D dan F dengan rentang nilai akurasi 1100 - 1500. Namun dapat dilihat model juga memiliki False Negative dan False Positive di kelas E, sehingga dapat dinyatakan bahwa model ini kurang akurat memprediksi kelas E.

1. Learning Curves



Dari hasil learning curve yang didapatkan nilai yang didapatkan itu rata rata bisa dilihat 0.950 dan ini merupakan hasil yang baik untuk data pelatihan tetapi di dalam validation learning curvenya itu mendapatkan accuracy yang lebih kecil dari training curve karena model mengalami kesulitan sehingga bisa terjadinya overfitting.