Na aplicação do algoritmo SVM na base utilizada a acurácia variou de 50% até 100% em alguns testes, com a base de testes selecionada em 20%. As SVM linear apresentaram na maioria das execuções uma acurácia de 75%, mas hora ou outra apresentava 50% e 100%, o mesmo para a rbf e sigmoid, no entanto para essas apresentando maioria da acurácia em 50%. Para uma base de testes de 30% a acurácia resultante também foi igualmente variável, mas nesse caso entre 33,33% e 66,66%, no entanto com todos os modelos apresentado resultados variáveis. Já com os dados balanceados, apresentou uma piora no senário, todos os modelos permaneceram igualmente variáveis entre 16% e 66,66%, muitas vezes apresentando acurácia de 50% e na maioria 33,33%.

Já na aplicação de arvores de decisão foram aplicados dois algoritmos, arvore de decisão comum com base na entropia dos dados e Random Forest que faz uso de várias arvores e mostra uma média dos resultados. Em relação aos retornos, com uma base balanceada e 30% dos dados para teste, os dois algoritmos tiveram uma acurácia variando entre 33% e 50%. Não houve melhoria no cenário mudando a quantidade de dados para teste. Já com a base desbalanceada e 20% dos dados de treinamento, foram apresentadas acurácias variáveis entre 25%, 50% e 75%. Com 30% dos dados para treino também resultou em acurácias variáveis para os dois algoritmos, 50%, 66,66%, 83,33% e até 100% em alguns testes. Em relação ao Pyplot da Arvore de decisão, pode ser observado que para essa base em relação ao tipo de dado analisado, pode deixar o usuário confuso, justamente cada classe apresentada na arvore indicar uma palavra que foi transformada pelo Bag-of-words.

Por fim é possível notar que os algoritmos de SVM e arvores de decisões, bem como suas variantes, apresentam resultados variáveis para base de teste em questão, mesmo balanceando base de treino e utilizando quantidades variáveis para validação. O que pode ser um indicio de que os mesmos algoritmos podem não ser tão eficientes para base em questão.

Aluno: Ivo Ireneu de Souza Juniors

Professor: Victor Lundgren

Disciplina: TAIA-SI-UAST