

Avaliação de Modelos

Graziele Fagundes Martins e João Vitor Farias

Avaliação dos Modelos e Resultados

Os modelos testados, Random Forest e AdaBoost, são algoritmos de aprendizado de máquina amplamente utilizados para tarefas de classificação. Ambos são baseados em conjuntos de árvores de decisão, mas possuem abordagens distintas para combinar as previsões dessas árvores. Abaixo, explicamos o funcionamento de cada um deles, bem como uma análise dos resultados obtidos nos datasets de Análise de Sentimentos e Previsão de Emojis.

1. Random Forest

O Random Forest (Floresta Aleatória) é um método de ensemble learning que combina múltiplas árvores de decisão para melhorar a precisão e reduzir o overfitting. O algoritmo funciona da seguinte forma:

- **Construção das árvores:** Cada árvore é treinada em um subconjunto aleatório dos dados de treinamento (amostragem com substituição, conhecida como bootstrap). Além disso, em cada divisão (split) da árvore, apenas um subconjunto aleatório de features é considerado. Isso introduz diversidade entre as árvores.
- **Previsão:** Para classificação, cada árvore "vota" na classe predita, e a classe com mais votos é escolhida como a previsão final.

O Random Forest é conhecido por sua robustez, capacidade de lidar com dados desbalanceados e resistência ao overfitting. No entanto, ele pode ser computacionalmente caro e menos interpretável do que uma única árvore de decisão.

2. AdaBoost

O AdaBoost (Adaptive Boosting) é outro método de ensemble learning, mas ele funciona de maneira diferente do Random Forest. O AdaBoost combina vários classificadores fracos (geralmente árvores de decisão simples, chamadas de "stumps") em um classificador forte. O algoritmo funciona da seguinte forma:

- **Treinamento iterativo:** O AdaBoost treina os classificadores fracos sequencialmente. Em cada iteração, ele dá mais peso às instâncias que foram classificadas incorretamente pelo classificador anterior, focando nos exemplos mais difíceis.
- **Previsão:** Cada classificador fraco contribui para a previsão final com um peso proporcional à sua precisão. A previsão final é uma combinação ponderada das previsões de todos os classificadores.

O AdaBoost é eficaz para melhorar a precisão de classificadores simples, mas pode ser sensível a dados ruidosos e outliers, pois ele tende a se concentrar em exemplos difíceis.

3. Análise dos Resultados

3.1. Dataset: Análise de Sentimentos ([Dataset](#))

- Acurácia do Random Forest: 0.528
- Acurácia do AdaBoost: 0.53

Ambos os modelos tiveram desempenho semelhante, com uma acurácia em torno de 53%. Isso sugere que a tarefa de análise de sentimentos pode ser desafiadora para esses modelos, possivelmente devido a:

- Complexidade do texto: A análise de sentimentos envolve nuances linguísticas, como sarcasmo, ironia e contextos específicos, que podem ser difíceis de capturar com modelos baseados em árvores.
- Qualidade dos dados: Se o dataset estiver desbalanceado ou contiver ruído, isso pode afetar o desempenho dos modelos.
- Limitações dos modelos: Random Forest e AdaBoost não são naturalmente adequados para processamento de linguagem natural (NLP), pois não capturam bem a semântica do texto. Modelos baseados em redes neurais, como LSTMs ou Transformers, podem ser mais adequados.

3.2. Dataset: Previsão de Emojis ([Dataset](#))

- Acurácia do Random Forest (Emoji): 0.247
- Acurácia do AdaBoost (Emoji): 0.221

Aqui, o desempenho dos modelos foi significativamente pior, com acurácias abaixo de 25%. Isso indica que a tarefa de prever emojis é ainda mais complexa do que a análise de sentimentos. Possíveis razões incluem:

- Alta dimensionalidade: A previsão de emojis pode envolver um grande número de classes (emojis), o que torna a tarefa mais difícil.
- Contexto cultural e linguístico: Emojis muitas vezes dependem de contextos específicos e nuances culturais, que podem não ser bem capturados por modelos baseados em árvores.
- Falta de features relevantes: Se o dataset não contiver features suficientes ou relevantes para a previsão de emojis, os modelos terão dificuldade em aprender padrões úteis.

4. Comparação entre os Modelos

Random Forest teve um desempenho ligeiramente melhor do que o AdaBoost na tarefa de previsão de emojis, mas ambos tiveram resultados ruins. Isso pode ser devido à natureza do AdaBoost, que tende a se concentrar em exemplos difíceis, mas pode não generalizar bem para tarefas complexas como a previsão de emojis.

Na análise de sentimentos, ambos os modelos tiveram desempenho semelhante, o que sugere que a tarefa pode estar no limite da capacidade desses algoritmos.

5. Possíveis Melhorias

Para melhorar o desempenho dos modelos, as seguintes abordagens podem ser consideradas:

1. Engenharia de features:
 - Criar features adicionais que capturem melhor o contexto do texto, como n-grams, polaridade de palavras, ou embeddings pré-treinados.
2. Modelos mais avançados:
 - Experimentar modelos baseados em redes neurais, como LSTMs, GRUs ou Transformers (e.g., BERT), que são mais adequados para tarefas de NLP.
 - Utilizar transfer learning com modelos pré-treinados em grandes corpora de texto.
3. Ajuste de hiperparâmetros:
 - Realizar uma busca sistemática pelos melhores hiperparâmetros para Random Forest e AdaBoost, como número de árvores, profundidade máxima e taxa de aprendizado.
4. Avaliação mais detalhada:
 - Além da acurácia, utilizar métricas como precisão, recall, F1-score e matriz de confusão para entender melhor o desempenho dos modelos, especialmente em datasets desbalanceados.

5. Conclusão

Os resultados obtidos com Random Forest e AdaBoost nos datasets de análise de sentimentos e previsão de emojis indicam que esses modelos têm limitações significativas para tarefas complexas de NLP. Embora sejam algoritmos poderosos para muitos problemas de classificação, eles não são naturalmente adequados para capturar a semântica e o contexto necessários para tarefas de linguagem natural.

Para melhorar o desempenho, é recomendável explorar modelos mais avançados, como redes neurais e técnicas de NLP modernas, além de investir em um pré-processamento cuidadoso dos dados e engenharia de features. Essas abordagens podem ajudar a extrair padrões mais significativos dos dados e, conseqüentemente, melhorar a acurácia e a generalização dos modelos.