Quality and Emotion Visualization in Text

Principais Caracteristicas:

- Base de Dados
- Algoritmos
- Heurísticas
- Modelo de Treino
- Aplicação Web

Base de Dados

- A base de dados contém 406 textos;
- Todos os textos têm mais de 5000 palavras e 60 frases.

Estrutura	Muito Simples	Pouco Complexo	Complexo	Muito Complexo
Nº Textos	8	238	138	20

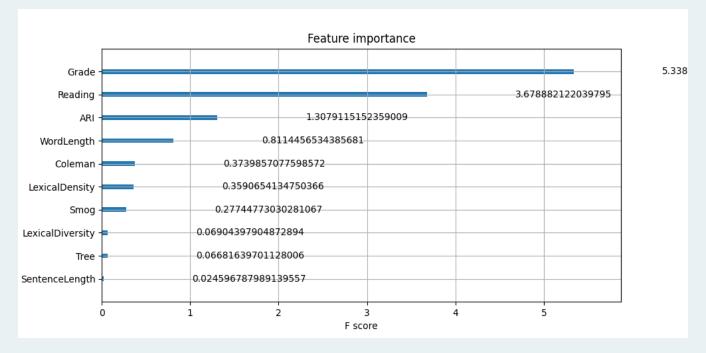
Sentimentos	Negativo	Neutro	Positivo
N° Textos	51	141	212

Amostras

- Tipo de Amostras: São retiradas 60 frases aleatórias do texto;
- Tipo de Frases: São apenas retiras frases que contem mais de 4 palavras;
- Razão: As amostras são retiradas desta forma para um funcionamento normal dos algoritmos desenvolvidos;
- Erro: Margem de erro comparado com a avaliação total do texto apenas de 1.2% a 10%.
- Eficiência: O uso de amostras demonstrou um aumento de performance de 90%.

Algoritmos

- Foram desenvolvidos 11 algoritmos no total;
- Medidas de Complexidade Textual: Diversidade e densidade lexical, tamanho médio das palavras e das frases e profundidade sintática da frase;
- Métricas de Leitura: SMOG, ARI, Coleman-Liau Index, Flesch Kincaid Grade Level e Flesch Reading Ease;
- Análise de Sentimentos é usado o módulo Spacy para essa avaliação.



Heurísticas

Estrutura Textual:

- Caracteristica comum: Quanto mais elevado o valor, melhor estruturado se encontra o texto;
- Aplicada normalização nos algoritmos para um resultado mais preciso e melhor manipulação dos dados;
- Aplicado a média de todos os algoritmos utilizados.

Visualização Sentimental:

É utilizado o compound como valor principal para avaliação sentimental do texto.

Heurísticas

Estrutura	Muito Simples	Pouco Complexo	Complexo	Muito Complexo
Intervalos:	[0, 0.25[[0.25, 0.50[[0.50, 0.75[[0.75, 1]

Sentimentos	Negativo	Neutro	Positivo
Intervalos:	[-1, -0.05[[-0.05, 0.05]]0.05, 1]

Modelo de Treino XGBoost

- Funções relevantes para o trabalho desenvolvido como: plot importance;
- Modelo mais preciso e eficiente comparado com modelos testados, árvore de decisão e regressão linear.

Modelos	Regressão Linear	Árvore de Decisão	XGBoost
Taxa de Acerto(%) Estrutura	73	93	96
Taxa de Acerto(%) Sentimento	43	98	99

Modelo de Treino XGBoost

- Usado para treino a percentagens comuns de 30% para testes e 70% para treino;
- A média dos testes foi obtida através da media feita a partir de 1000 treinos diferentes.

Classificadores	F1 Score	Recall	Taxa de Acerto	Precisão
Percentagem(%) Estrutura	96.02	96.11	96.11	96.27
Percentagem(%) Sentimento	99.1	99.11	99.1	99.13

Aplicação Web

Quality and Emotion Visualization in Text

Resultado de Qualidade Textual

O texto analisado obteve uma classificação de: **Estrutura Pouco Complexa**

Estrutura Muito Complexa: Teor profissional, estrutura complexa e leitura mais difícil.

Estrutura Complexa: Estrutura bem desenvolvida e leitura de nível médio.

Estrutura Pouco Complexa: Estrutura menos desenvolvida e leitura fácil.

Estrutura Muito Simples: Estrutura fraca, baixa complexidade e leitura extremamente fácil.

143%

Resultado da Análise dos Sentimentos

O texto analisado obteve um resultado sentimental: Negativo

Positivo: O texto contém sentimentos agradáveis, como alegria, prazer ou entusiasmo.

Neutro: O texto não contém sentimentos relevantes, sendo provavelmente informativo, como uma notícia ou artigo.

Negativo: O texto contém sentimentos desagradáveis, como tristeza, raiva ou ansiedade.

The rain poured down relentlessly, each drop a small, cold reminder of the emptiness that now filled his heart. The room, once lively and full of laughter, stood in stark silence, echoing with memories that seemed too painful to bear. The photographs on the walls, once treasured moments captured in time, now felt like haunting specters of a past that could never be reclaimed. Every corner of the house whispered her name, her absence a gaping void that no amount of tears could fill. He wandered from room to room, searching for traces of her, a scent, a forgotten piece of clothing, anything that would make her feel close again. But all he found were shadows and silence, a chilling reminder that she was truly gone.

The world outside carried on, indifferent to his pain. The sun rose and set, people laughed and lived, while his world had come to a devastating halt. The bed they once shared felt like a grave, the cold sheets a stark contrast to the warmth she had once brought into his life. Nights were the worst, he lonelines amplified by the darkness, his cries swallowed by the emptiness around him. He remembered the day she left, the way her eyes had looked, filled with sadness and resignation. They had fought, harsh words thrown in moments of anger and frustration, words he wished he could take back. She had walked out, leaving behind a silence that spoke louder than any argument ever could. He had thought she would come back, that they would mend the broken pieces, but days turned into weeks, and weeks into months, with no sign of her return.

Her laughter, once a melody that filled their home, now haunted his dreams. He saw her everywhere, in the crowds of strangers, in the shadows of their favorite places. He reached out, only to grasp at the cold, cruel air. Friends tried to comfort him, their words well-meaning but hollow, unable to penetrate the depths of his despair. They told him it would get better, that time would heal, but time only seemed to deepen the wound. Her side of the closet remained untouched, clothes still hanging as if she would walk in any moment and pick out her favorite dress. He couldn't bring himself to move them, to erase the last remnants of her presence. Her scent lingered faintly, a cruel reminder of what he had lost. Each morning he woke with a heavy heart, the weight of her absence pressing down on him, making it hard to breathe.

He tried to distract himself, throwing himself into work, into meaningless tasks, but nothing could fill the void. Nights were spent staring at the ceiling, replaying every moment, every mistake, wishing he could turn back time. He remembered their last holiday, the way she had smiled, the way her eyes had sparkled with joy. That memory now fell like a distant dream, a cruel reminder *

Algoritmos Usados:

Reading com uma importância de: 5.858
Grade com uma importância de: 2.305
ARI com uma importância de: 1.737
LexicalDensity com uma importância de: 0.797
Coleman com uma importância de: 0.586
WordLength com uma importância de: 0.528
Smog com uma importância de: 0.369
Tree com uma importância de: 0.154
LexicalDiversity com uma importância de: 0.151
SentenceLength com uma importância de: 0.099

Texto Exemplos

Texto 2

Texto 3

Analisar Texto

Resultados dos Algoritmos: ARI-0.607 Coleman-0.684 Grade-0.432 LexicalDensity-0.49 LexicalDiversity-0.43 Reading-0.276 SetnenceLength-0.152 Smog-0.777 Tree-0.399 WordLength-0.385

Aplicação Web

Resultado de Oualidade Textual

O texto analisado obteve uma classificação de: Estrutura Pouco Complexa

Estrutura Muito Complexa: Teor profissional, estrutura complexa e leitura mais difícil.

Estrutura Complexa: Estrutura bem desenvolvida e leitura de nível médio.

Estrutura Pouco Complexa: Estrutura menos desenvolvida e leitura fácil.

Estrutura Muito Simples: Estrutura fraca, baixa complexidade e leitura extremamente fácil.

36%

Resultado da Análise dos Sentimentos

O texto analisado obteve um resultado sentimental: Negativo

Positivo: O texto contém sentimentos agradáveis, como alegria, prazer ou entusiasmo.

Neutro: O texto não contém sentimentos relevantes, sendo provavelmente informativo, como uma notícia ou artigo.

Negativo: O texto contém sentimentos desagradáveis, como tristeza, raiva ou ansiedade.

Zona responsável por mostrar de forma listada os algoritmos utilizados

Zona responsável por mostras os resultados obtidos e explica-los

Algoritmos Usados:

Reading com uma importância de: 5.858

Grade com uma importância de: 2.305

ARI com uma importância de: 1.737

LexicalDensity com uma importância de: 0.797

Coleman com uma importância de: 0.586

WordLength com uma importância de: 0.528

Smog com uma importância de: 0.369

Tree com uma importância de: 0.154

LexicalDiversity com uma importância de: 0.151

SentenceLength com uma importância de: 0.099

Zona responsável por mostrar os resultados normalizados dos algoritmos

Conclusões Finais

- Importância de bons algoritmos de avaliação estrutural e sentimental;
- Importância de boas heurísticas;
- Importância de um bom modelo de treino;
- Existência de uma relação entre todos os algoritmos usados.

Obrigado