



Análise e previsão de cores utilizando texto de cartas de Magic The Gathering

Text Mining

Docente: Ana Martins

David Isaac, N 120064

Eduardo Martins, N° 120063

João Figueiredo, N° 120046

Dezembro 2024

Índice

Resumo	3
Introdução	4
1. Definição do problema	6
2. Compreensão de dados	9
3. Preparação de dados	11
4. Visualização	14
5. Modelação	20
Conclusão	26
Referencias bibliográficas	27

Resumo

Em Magic: The Gathering, jogo de cartas colecionáveis lançado em 1993 pela Wizards of The Coast, é jogado por um ou mais jogadores que têm como objetivo ser o último jogador restante na partida. Cada carta pode pertencer a uma ou mais de 5 cores distintas, sendo elas azul, vermelho, branco, preto e verde, e cada cor possui características únicas, tanto nos efeitos que produz em jogo, quanto no contexto em que ela está inserida na história por detrás de cada coleção lançada, esta característica do jogo foi apelidada pela comunidade como **filosofia das cores**.

Este projeto visa prever a cor das cartas através dos seus elementos textuais, nomeadamente o flavor text e o texto das habilidades. Os textos foram tratados de forma a não existirem referências diretas à cor correta, mas fornecem pistas sobre a caracterização das personagens e dos efeitos associados. O objetivo das modelações será buscar entender se a filosofia das cores está a ser devidamente aplicada na idealização das cartas, a observar tanto o efeito prático da carta em jogo, quanto na parte da sua história.

Para a realização desta previsão, o texto foi pre-processado e manipulado, nomeadamente através da tokenização, remoção de stopwords, conversão do texto para minúsculas e remoção de palavras que não se encontram no léxico inglês (como “aaaahhh” e “ac”).

Adicionalmente, através da utilização de regex, foram geradas novas features, provenientes dos elementos textuais que caracterizam as ações executadas pela carta ao ser utilizada durante o jogo. Estas features são *keywords* (Magic, n.d.), que são elas efeitos básicos, comuns à várias cartas, mas que tendem a concentrar-se em cores específicas, e outros efeitos que devem ser transversais a mais de uma carta de cada cor. Tais features foram então utilizadas para visualizações que mostraram ao grupo que de facto a **filosofia das cores** permanece presente na escolha dos efeitos que cada carta teria.

Por fim, foi utilizado o TF-IDF (Term Frequency-Inverse Document Frequency) para converter os textos do flavor e das ability em representações numéricas, de forma a permitir, a criação de modelos de Regressão Logística, Naive Bayes Multinomial e uma Rede Neuronal, escolhidos pela capacidade dos mesmos em avaliar e classificar textos.

Os resultados mostram que o modelo final alcançou uma precisão de 69%. Estes resultados permitem-nos observar que os elementos textuais podem fornecer informações suficientes para a previsão das cores das cartas, embora com algumas limitações em certas categorias.

Estes resultados nos permitem retirar informações valiosas, relativamente ao processo de criação do universo deste jogo e integração destes elementos narrativos nas cartas. O grupo acredita ainda ter encontrado uma possível falha pela equipa responsável por este processo e estuda quais decisões foram tomadas posteriormente pela fabricante para contornar isto.

Commented [JF1]: tive um derrame sobre como citar isso

Commented [JF2]: Gostaria que este termo fosse usado ao longo do relatório, para isso eu acho que ele deveria ser citado rapidamente aqui no resumo e elaborado bem mais na introdução, qual vossa opinião?

Commented [EM3R2]: Ok, eu nem conhecia esse termo, tem de se modificar na introdução então, mas tudo bem

Commented [JF4R2]: não é muito "oficial" mas assim temos uma determinação pra esse conceito que podemos usar ao longo do trabalho sem ter que estar sempre a explicar

Commented [DM5]: Que acham?

Commented [EM6R5]: Está bom

Introdução

O jogo Magic: The Gathering é um dos mais antigos e populares jogos de Trading Card Games (TCG), com uma vasta coleção de cartas impressas desde o seu lançamento em 1993. O jogo foi criado por Richard Garfield, lançado pela empresa Wizards of the Coast, sendo esta posteriormente adquirida pela Hasbro como subsidiária. No jogo, cada carta é desenhada com o intuito de pertencer a uma ou mais cores, que representam temáticas específicas no jogo. Por exemplo, cartas verdes geralmente apresentam características mais relacionadas com a natureza, crescimento e mecânicas de combate. Por outro lado, as cartas azuis estão mais associadas ao controlo e a manipulação.



Jogadores a competir em um torneio de Magic, retirado de The New York Times em dezembro de 2024

Cada carta tem vários elementos, como por exemplo, imagem, custos e raridade. Entre esses elementos existem dois que se destacam para a elaboração deste projeto: Flavor e ability Text. Estes dois correspondem a descrições narrativas e às mecânicas das cartas respetivamente e que não possuem qualquer tipo de referência direta à cor da carta. No entanto, estes fornecem pistas indiretas sobre a identidade temática da carta, como por exemplo referências a habilidades como “flying”, “fight” e “counter” que muitas vezes estão associadas a cores específicas.

Este projeto tenta estudar este desafio através do desenvolvimento de modelos que prevêm a cor das cartas de MTG utilizando somente a informação textual disponível de cada uma.

O dataset utilizado para a elaboração deste estudo, proveniente do Kaggle (Gendotti, 2022), contém a descrição exhaustiva de todos os elementos de todas as cartas impressas entre 5 de Agosto de 1993 e 20 de Maio de 2022, à exceção das suas imagens.

De forma a atingir este objetivo, vários passos foram necessários:

Commented [DM7]: H?

Commented [JF8]: Gente, atenção que citações no texto é só ultimo nome

Commented [DM9R8]: okidoki

- Pre-processamento de dados: O texto foi limpo e tokenizado, com stopwords e palavras não *standard* no léxico inglês de forma a aumentar a qualidade dos dados.
- Feature Engineering: Expressões Regulares (Regex) foram utilizadas para identificar mecânicas específicas que estão tematicamente associadas a cores de cartas específicas.
- Representação textual: Foi utilizado TF-IDF (Term Frequency-Inverse Document Frequency) de forma a captar a importância das palavras.
- Modelação: Modelos Preditivos foram utilizados, incluindo modelos mais tradicionais, nomeadamente Regressão Logística, Naive Bayes e Redes Neurais para classificar as cores das cartas.

A partir da análise dos resultados dos modelos, busca-se entender se, no geral, as cartas estão alinhadas com a sua cor, tanto nos seus efeitos, quanto em sua caracterização, ou se este fator importante para a ambientação do jogo não está a ser seguido devidamente. Estes insights podem ser relevantes para a equipe de criação entender quais pontos devem ser melhorados e quais são as melhores decisões a serem tomadas para que o jogo não perca esta característica importante, mantendo seus jogadores interessados.

1. Definição do problema

Magic The Gathering é um TCG, acrônimo para *Trading Card Game*, um jogo de cartas onde o objetivo principal é que os jogadores colecionem e troquem cartas entre si para criarem baralhos e competirem entre si em jogos contra um ou mais oponentes em formatos variados. O **Magic** por sua vez gira suas principais mecânicas em torno de cinco cores principais, cada uma possui estilos de jogo e estratégias diferentes das outras, esse conceito é batizado de **filosofia das cores** (Sabien, 2024).



Ilustração que resume a **filosofia das cores**

O vermelho foca-se na ação e na liberdade, querendo expressar a sua paixão e energia através da rapidez e agressividade. No jogo, isto traduz-se em dano direto, criaturas rápidas e poderosas, e efeitos que aceleram o ritmo do jogo. O vermelho prioriza ataques rápidos e fortes, com uma abordagem impetuosa e imprevisível, visando encurtar as partidas e provocar grandes explosões de poder. Busca a emoção, momentos intensos, impulsivos e apaixonantes durante as suas escolhas. Mas, quando em excesso, o vermelho pode se tornar descontrolado, sem estrutura e guiado apenas pela necessidade de seguir sempre em frente.

O azul, por sua vez, foca no controle e na manipulação, querendo alcançar a perfeição através do conhecimento e da estratégia. No jogo, isso traduz-se em anular as cartas adversárias, como contrafeitiços, e manipular o ritmo do jogo por meio de truque ilusórios e ganho gradual de vantagem. O azul concentra-se na clareza e na precisão, sempre em busca da verdade e da solução perfeita para qualquer problema.

O verde aposta na força e no crescimento, onde através da harmonia com a natureza, busca acelerar o desenvolvimento de recursos e invocar criaturas imponentes para dominar o campo de batalha. Ao abraçar a sabedora ancestral e o fluxo natural das coisas, o verde utiliza o poder da regeneração e da força bruta para restaurar o equilíbrio, crescendo em poder a medida que a batalha avança

O branco enfatiza ordem e sinergia, equilibrando o campo com defesas robustas e fortalecendo aliados com a força da cooperação e proteção mútua. Ao estabelecer regras claras e estruturas firmes, o branco cria um ambiente de estabilidade onde a união e a disciplina garantem a segurança e o sucesso coletivo

Já o preto explora riscos e sacrifícios, utilizando vida e criaturas como recursos para obter vantagens, enfraquecer o oponente e assumir o controle da partida. Ao se desprender de limitações morais ou éticas, o preto age com uma visão pragmática, disposto a pagar o preço necessário para atingir os seus objetivos.



Na esquerda *Dismiss*, um exemplo de carta azul que combina anulação com compra de cartas, e na direita *Lightning Bolt*, uma das maneiras que o vermelho possui de causar dano direto ao oponente e suas criaturas, retirado de *Card Market* em dezembro de 2024

As cartas são lançadas em coleções periódicas, que adicionam novas cartas e mecânicas ao jogo. Estas coleções possuem sua própria ambientação e as cartas representam criaturas e personagens daquele mundo, assim como pontos-chave da história que estas personagens estão a viver naquele mundo.



Na esquerda a personagem Gideon e na direita *Oath of Gideon*, ambos pertencentes à coleção *Battle for Zendikar* (2016), retirado de *Card Market* em dezembro de 2024

Gideon é uma personagem emblemática que já apareceu em diversas cartas ao longo dos anos, é um guerreiro justo que luta por aqueles que não podem e guia legiões de guerreiros ao combate, essa personalidade justa e capacidade estratégica em lutas são características da cor branca, tal como referido anteriormente, que irá representar a personagem em suas cartas e habilidades.

Ele aparece em *Battle for Zendikar*, coleção que tem como tema a luta entre diversos heróis, dentre eles Gideon, contra criaturas que buscam invadir o país de Zendikar. Em determinado momento da história, os heróis fazem o juramento de defender aquele lugar e muitos outros que aquelas criaturas estavam a atacar. Cada um dos juramentos foi representado por cartas, no exemplo acima está o de Gideon, que, para além do seu efeito, possui o texto adicional (*flavour text*) que regista a fala da personagem na cena representada. O texto mostra o desejo da personagem de manter a justiça, valor que está alinhado com branco segundo a **filosofia das cores**.

O objetivo será criar um modelo que faça a previsão de qual cor a carta pertence baseado nestes elementos textuais presentes nelas.

É bom notar que há cartas que possuem mais de uma cor, sendo que estas combinações possuem também suas próprias características, que geralmente são uma combinação das características das cores que a compõem. Para a abordagem deste trabalho, apenas as cartas com uma cor serão selecionadas, a fim de simplificar o problema.



Niv-Mizzet é um exemplo de personagem com mais de uma cor, mesclando a capacidade do vermelho de causar dano direto com a compra de cartas do azul, retirado de *Card Market* em dezembro de 2024

2. Compreensão de dados

O dataset usado neste estudo é originário do Kaggle e possui informações detalhadas acerca de todas as cartas entre 5 de agosto de 1993 e 20 de maio de 2022. O conjunto de dados abrange uma ampla variedade de atributos relacionados às cartas, fornecendo um panorama completo das suas características.

No dataset original estão presentes 39 colunas, representando diferentes informações sobre as cartas, desde o nome e o texto de habilidades até a características mais específicas como a cor e o tipo de mecânicas associadas. No entanto, nem todas elas são relevante para o nosso estudo, que consiste na previsão das cores das cartas com base nos seus elementos textuais. Assim, foi necessário realizar uma seleção de features para garantir que apenas as colunas mais pertinentes fossem utilizadas na análise.

As seguintes features foram consideradas relevantes para o estudo:

- ‘name’: Esta feature corresponde ao nome da carta. Embora não seja utilizada diretamente para prever a cor, pode auxiliar na identificação de cartas únicas durante a análise.
- ‘color_identity’: Indica a cor associada a cada carta, podendo ser White (W), Blue (U), Black (B), Red (R) ou Green (G). Para simplificar o problema de classificação, foram consideradas somente cartas com uma única cor. Cartas com múltiplas cores

foram removidas, reduzindo assim a complexidade da tarefa e focando na relação direta entre texto e cor.

- ‘text’: Esta feature corresponde ao texto de habilidade presente em cada carta. Este texto descreve os efeitos e mecânicas que podem ser utilizados em combate. A importância desta feature reside no facto de que existem habilidades que estão normalmente associadas a cores específicas. Por exemplo, *counterspells* estão frequentemente associadas à cor azul.
- ‘flavor’: Esta feature corresponde a um texto narrativo ou descritivo que não está ligado às habilidades da personagem, mas sim com a sua caracterização em si. Esta feature é relevante uma vez que certas espécies (e.g. dragões) e traços de personalidade (e.g. “errático”) estão associadas a cores específicas.

3. Preparação de dados

Inicialmente, com o intuito de mais tarde se efetuarem visualizações que nos permitissem ter uma melhor visualização das possíveis relações existentes entre cores e habilidades, foram criadas colunas booleanas, através da aplicação de expressões regulares (regex) no texto de habilidades, que nos permitissem agrupar as diferentes cartas pelas suas habilidades. Por exemplo, a habilidade “ganhar vida” a coluna foi criada da seguinte forma: `df["ganhar_vida"] = (df.text.str.contains(r"gain(?:\s|\s\s)\d+\s\slife"))`. Deste modo, todas as cartas com a habilidade “ganhar vida” teriam o valor True nesta coluna. Este processo foi repetido para outras habilidades relevantes existentes no jogo.

Seguidamente, foram removidas informações que se encontrassem dentro de parênteses. Essas informações, conhecidas como texto de lembrete, servem para esclarecer como certas mecânicas de jogo funcionam e não estão diretamente associadas ao texto das cartas. Estas informações podem incluir referências a outras mecânicas, as quais podem introduzir ruído para a análise.

Posteriormente, foi removida toda a pontuação das frases, garantindo assim que apenas caracteres alfabéticos são preservados. Foram também removidas palavras do texto que continham números, caracteres especiais ou letras excessivamente repetidas (e.g. onomatopeias), deixando apenas palavras válidas no léxico inglês.

Foram também substituídos todos os valores nulos por uma string vazia e foi adicionada uma colunas que agregasse o texto ao flavor.

De forma a prevenir que palavras muito semelhantes fossem contempladas múltiplas vezes (e.g. “goblin”, “goblins”), lematizou-se todos os tokens.

Mais tarde, aplicou-se TF-IDF com o objetivo de transformar os textos das cartas em valores numéricos, com os quais seja possível trabalhar. Esta abordagem captura a importância de cada palavra no contexto textual da carta, atendendo à frequência relativa no conjunto total de cartas. Esta abordagem permite ainda reduzir o peso de palavras muito comuns, destacando palavras mais relevantes para a descrição de cores de forma única.

Ao analisar o top 20 de palavras com mais “impacto”, obtemos a seguinte tabela:

	color	word	tfidf
85613	W	creature	0.606746
24131	G	creature	0.592940
3637	B	creature	0.572885
44625	R	creature	0.496761
63842	U	card	0.430140
65119	U	creature	0.413124
78994	U	target	0.407452
58500	R	target	0.375678
2360	B	card	0.355982
44896	R	damage	0.345607
99488	W	target	0.338763
17512	B	target	0.338143
45044	R	deal	0.311929
85324	W	control	0.296473
59587	R	turn	0.280656
30006	G	target	0.270255
23842	G	control	0.263582
22054	G	card	0.255479
180575	W	turn	0.253293
39093	G	turn	0.244839

Top 20 palavras com maior "impacto" e as suas cores correspondentes

Commented [JF10]: Mudei aqui pq n foram todas nem de perto
Tem mm bués, algumas temporárias (aparecem em poucos sets) eu ignorei

Commented [JF11]: Adicionar alteração posterior do Isaac que retira novas palavras

Assim, concluiu-se que existem palavras que aparecem em todas as classes e, como tal, não apresentam cariz distinto o suficiente para contribuírem significativamente para a classificação das cores.

Em conjunto com posteriores análises de previsões prematuras, conseguimos, através de gráficos de barras e world clouds, identificar mais palavras que: apareciam em grandes quantidades, em várias cores; provocavam escolhas erradas de cores através de ruído.

De forma a ultrapassar isto, decidiu-se remover-se estas palavras, isto é, as palavras cuja relevância está presente em todas as cores (como as StopWord), e as palavras cujo a sua presença provocava ruído no modelo.

Deste modo, o método utilizado para a remoção destas palavras foi o seguinte: Inicialmente, calcula-se o valor máximo de importância de TF-IDF associado à palavra. De seguida seleciona-se um valor proporcional ao valor máximo, neste caso, o valor selecionado foi $\frac{1}{3}$ do valor máximo. Finalmente, verifica-se se as importâncias das palavras nas outras cores (com a exceção da cor com valor mínimo) se encontram dentro desse intervalo. Se a importância da palavra é próxima em múltiplas cores, esta é considerada StopWord e adicionada a uma lista de palavras a serem removidas mais tarde. Esta remoção afeta tanto o conjunto de dados TF-IDF como o texto processado no conjunto de treino.

Posteriormente, em relação as palavras que provocavam ruído no algoritmo, decidimos dar “append” das mesmas, na lista das palavras a serem removidas.

Commented [EM12R11]: @João Vitor Figueiredo Assim?

Commented [JF13R11]: falta explicar de forma mais detalhada como é feito a seguir a essa introdução

Commented [EM14R11]: Mas pensei que o isaac tivesse dito que na preparação de dados era só um cheirinho e na modelação fosse mais a fundo. Mas posso aprofundar mais se quiserem aqui

Commented [JF15R11]: hmmm acho q é aqui

Commented [JF16R11]: eu embaixo menciono que nós fizemos esse processo após reparar nas repetições, então faz sentido eu acho

Commented [EM17R11]: @João Vitor Figueiredo, vê se é mais isto que querias

Commented [JF18R11]: bué bom

Commented [EM19R11]: Ótimo, puta de dor de cabeça, n tens noção hahaha

Commented [DM20]: Nova informação

Commented [DM21]: Nova informação

Commented [JF22R21]: houve melhora? pq algumas palavras que tu retirou tipo graveyard me pareceram estranhas
preto mexe muito com cemitério

Commented [DM23R21]: Não melhorou nem piorou, tecnicamente era por culpa do branco

Commented [DM24R21]: Mas eu volto a meter para ver again

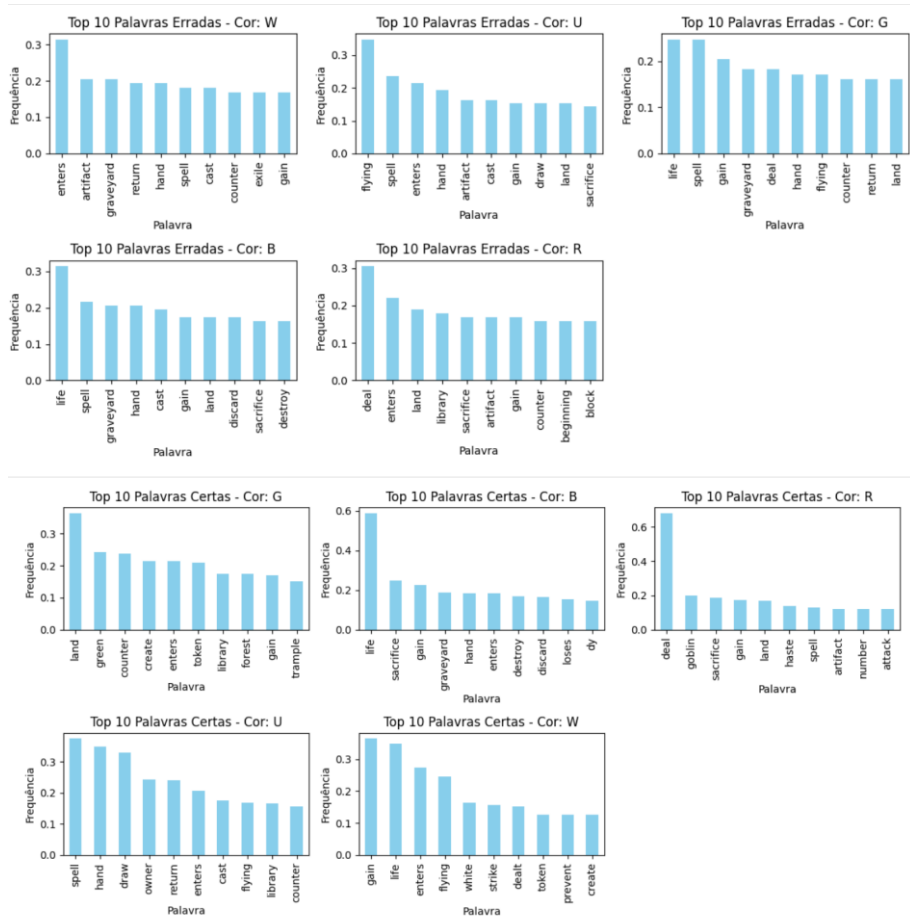


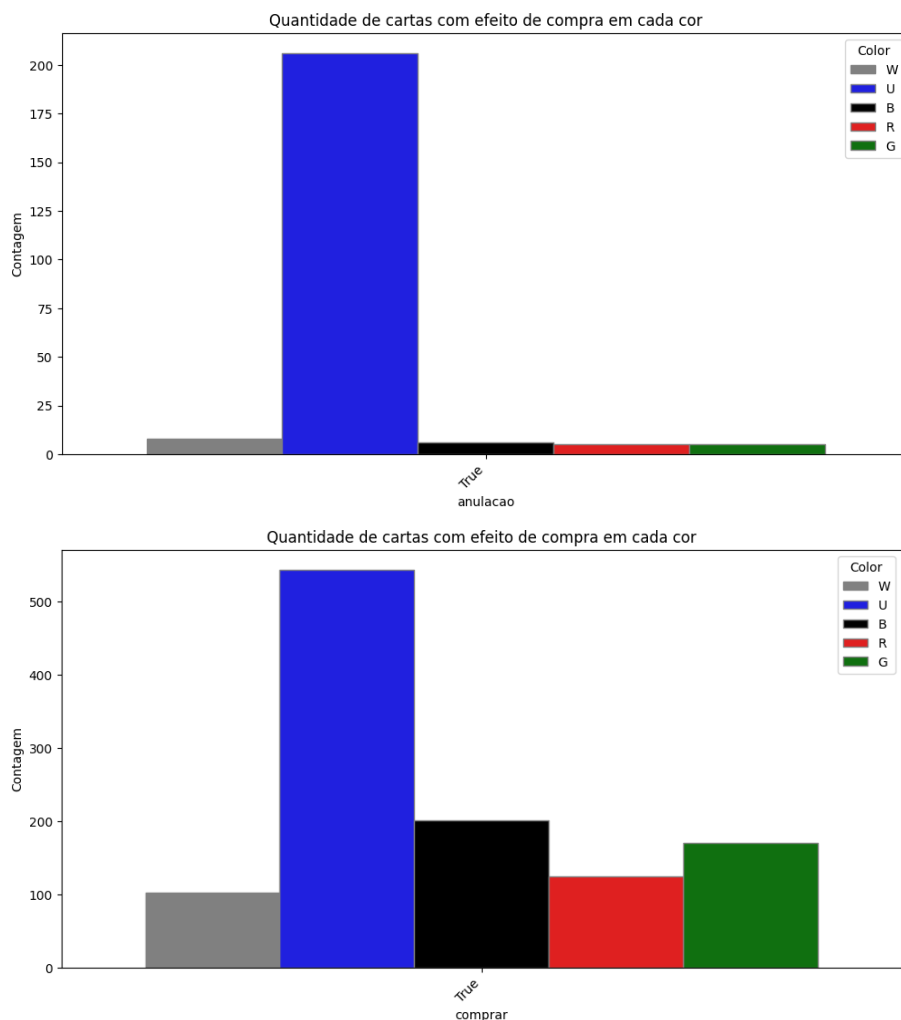
Gráfico de barras com a frequência palavras certas e erradas aparecem nas varias cores das cartas

4. Visualização

Nesta seção, buscaremos visualizações que mostrem que formas a **filosofia das cores** está presente no texto das cartas. Este capítulo irá revisar características citadas anteriormente para cada cor a fim de observar se elas estão de facto distribuídas da maneira proposta.

Para uma análise inicial do texto de regra, foi utilizado o dataset com palavras-chave, obtido através do uso de *Regex*. Com ele vamos observar se a presença de certas palavras-chave está alinhada com o que foi estudado acerca de cada cor.

Para o azul, que deseja tornar o jogo mais lento e controlado, é esperado uma maior presença de anulações e ferramentas de compra quando comparado com outras cores.

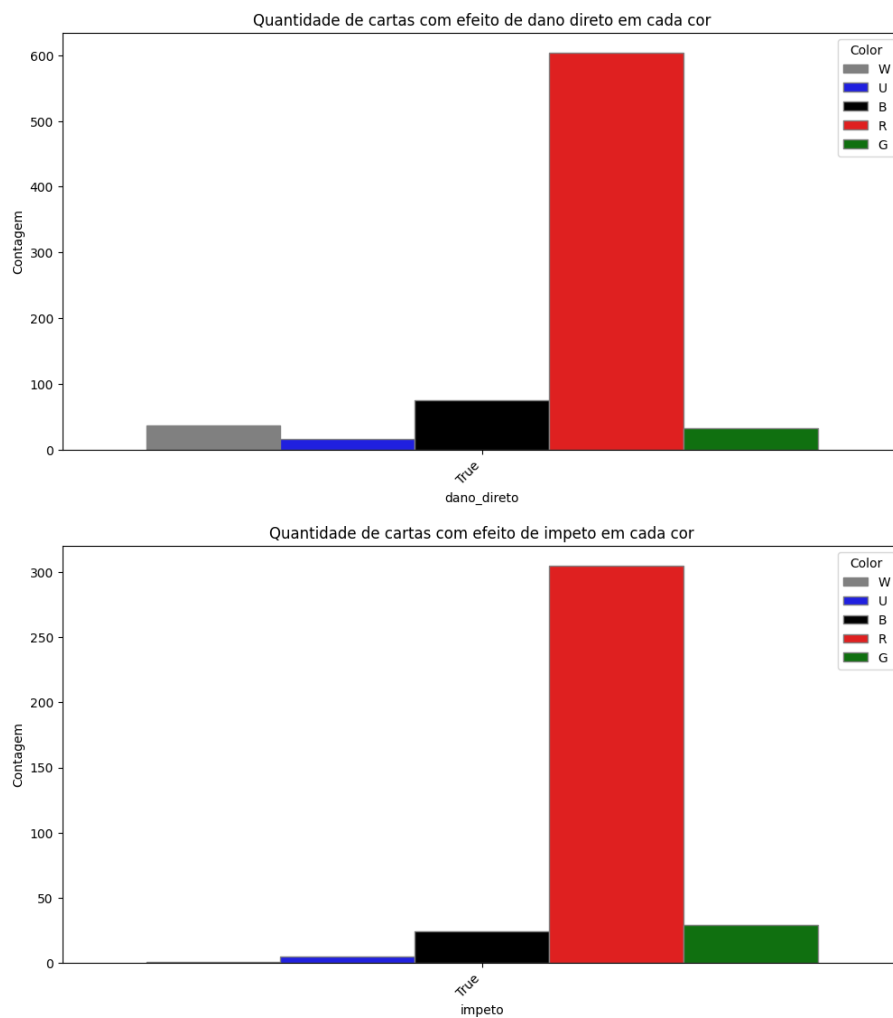


Podemos perceber que o efeito de anulação é praticamente exclusivo do **Azul**. Já a compra de cartas, embora mais predominante no azul, também está presente em outras cores. Ao analisar mais a fundo, percebe-se, no entanto, que a compra de cartas nas outras cores vem geralmente com um custo maior e associada com outras características destas cores. Por exemplo, compras de cartas verdes geralmente irão precisar de criaturas em jogo para funcionar, enquanto o preto costuma adicionar como custo o sacrifício de outras criaturas ou o pagamento de vida do jogador.



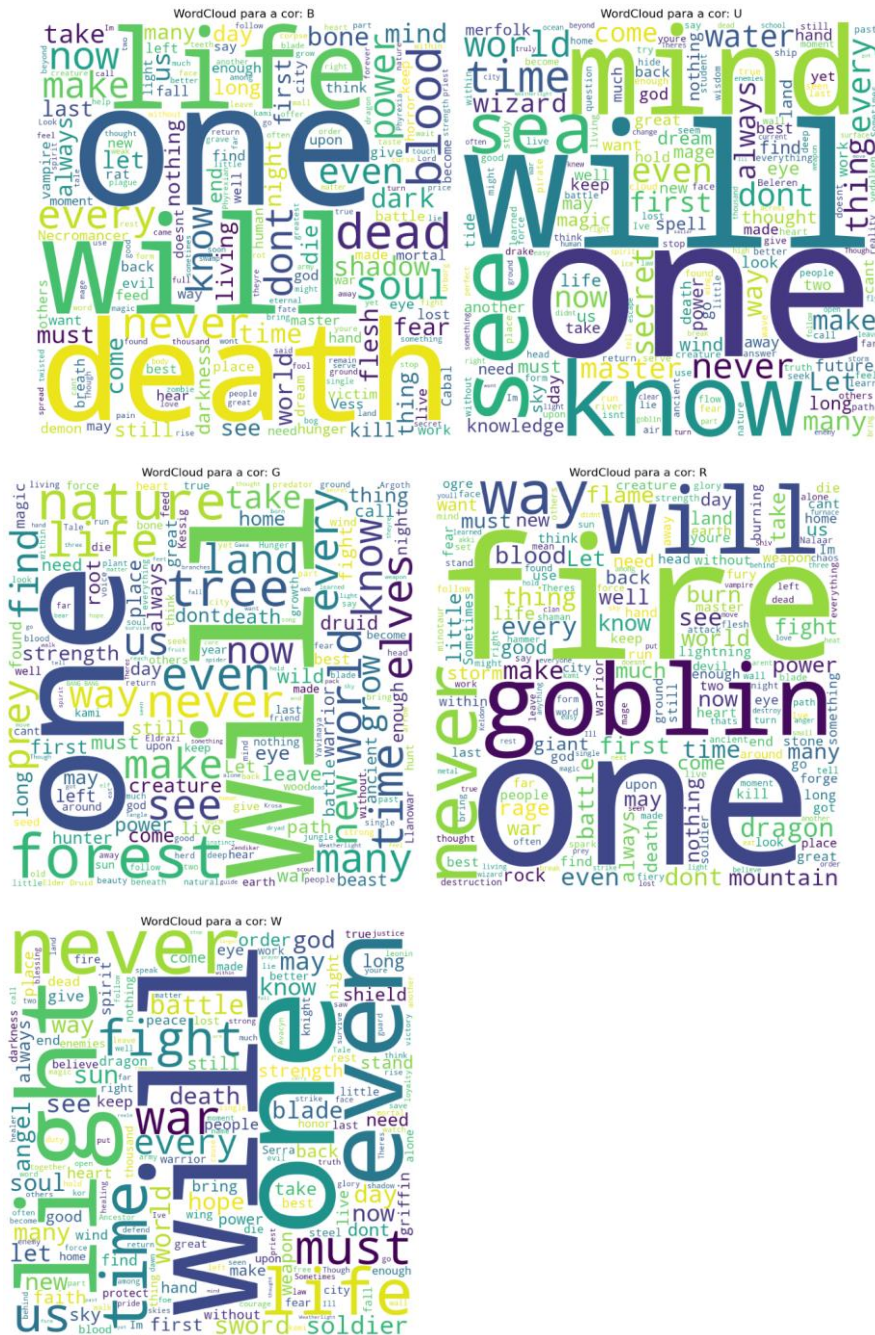
Exemplo de cartas não-azuis que compram cartas com condições alinhadas à sua cor, ambas disponibilizadas na coleção *Thrones of Eldraine* (2019), retirado de *Card Market* em dezembro de 2024

Passando para o vermelho, devido ao seu objetivo de colocar dano na mesa de maneira rápida, possui cartas com acesso a dano direto e criaturas com ímpeto (ou *haste* em inglês), o que permite que estas ataquem de imediato ao serem jogadas. Isso pode ser observado nos gráficos a seguir.



Concluimos assim, que a coluna Text pode conter informação útil para ser usada para treinar os modelos.

Já para o *flavour text*, o grupo utilizou *wordclouds* para encontrar palavras recorrentes em cada cor.



Wordclouds gerados com base na frequência das palavras em cada cor

O grupo identificou que existem palavras que se repetem bastante em todas as cores, dentre elas "*one*" e "*will*" são as mais aparentes, estando entre as palavras mais predominantes em todas as categorias. Motivado por esta análise, a função de detecção e retirada de *stopwords* foi mais bem trabalhada, a fim de detectar e retirar palavras com importância similar em múltiplas cores, tal como descrito no capítulo anterior.

Mesmo assim, alguns insights sobre a natureza das cores na história do jogo podem ser retirados da visualização.

Foi visto que o preto possui mecânicas relacionadas a sacrifícios, tanto de criaturas quanto de vida, e utiliza o cemitério como recurso. Dessa maneira, vemos palavras como "*death*", "*life*", "*soul*" e "*blood*", que remetem a essas características de vida e morte como um recurso. Algumas criaturas que se relacionam com o obscuro como vampiros, *zombies* e horrores também são mencionados.

A natureza caótica do vermelho é representada por diferentes referências a fogo e chamas, assim como sangue e tempestades. Criaturas como gigantes, dragões e principalmente goblins, seres gananciosos e impetuosos, presentes em literaturas de fantasia, marcam essa cor.

O controle do azul é representado por referências à mente e conhecimento, assim como água e mar. Inspirados em sereias, os "*merfolks*" aparecem devido às suas ligações ao mar e à água. Por fim, existem também menções à magia e conjuradores, "*wizard*" e "*mage*".

No verde há algumas alusões à natureza, como "*forest*", "*nature*" e "*tree*". Entre as criaturas presentes estão elfos e druidas, entre outras referências mais genéricas a outros seres como predador e presa.

Por fim, no branco há imensa presença da palavra luz, embora ela também estivesse presente em outras cores, e outras que remetem à ideia de manter a ordem, "*Justice*", "*peace*" e "*protect*" são alguns exemplos. As criaturas brancas mais emblemáticas que aparecem no wordcloud são os anjos e soldados.

Os resultados destas visualizações mostram que características únicas a cada uma das cores tendem a existir nestes textos. Sendo assim, a coluna *flavor* pode ser relevante durante a criação do modelo preditivo.

5. Modelação

No capítulo a seguir estão registados os resultados dos modelos para cada uma das abordagens escolhidas. Os modelos testados foram *Naive Bayes* e *Logistic Regression*, ambas em conjunto com TF-IDF, e foram treinados com as colunas *flavor* e *text* individualmente e posteriormente com a junção das duas. Buscamos aqui avaliar seus resultados, a fim concluir se a equipe de desenvolvimento do jogo continua alinhada com **filosofia das cores**, tanto no que diz respeito às regras, tanto à história contada através das coleções.

A primeira abordagem testada foi utilizando o *flavour text*. Os resultados se encontram abaixo.

```
Model: TF-IDF + Naive Bayes
Accuracy: 0.49
Model: TF-IDF + Logistic Regression
Accuracy: 0.48
```

Naive Bayes metrics:

	precision	recall	f1-score	support
R	0.47	0.44	0.45	489
U	0.53	0.50	0.51	488
G	0.48	0.48	0.48	518
W	0.55	0.42	0.48	511
B	0.46	0.60	0.52	531
accuracy			0.49	2537
macro avg	0.50	0.49	0.49	2537
weighted avg	0.49	0.49	0.49	2537

Logistic metrics:

	precision	recall	f1-score	support
R	0.43	0.45	0.44	489
U	0.53	0.46	0.49	488
G	0.51	0.47	0.49	518
W	0.49	0.46	0.48	511
B	0.46	0.56	0.50	531
accuracy			0.48	2537
macro avg	0.48	0.48	0.48	2537
weighted avg	0.48	0.48	0.48	2537

Os resultados foram abaixo do esperado, com menos de 50% de precisão. Ambos os modelos tendem a acertar mais cartas azuis, com o Naive Bayes a acertar também mais branco, enquanto a Logistic Regression acerta mais verdes, em comparação com as restantes cores.

Conclui-se que, com base nos resultados desanimadores, a equipe responsável pela criação da história e ambientações do jogo não está a ser capaz de tornar notável estas características da **filosofia das cores** naquilo que é publicado. Pesquisando mais a fundo, percebe-se que esta parte de construção de histórias e mundos foi marcada por insucessos.

O primeiro deles foi o encerramento dos servidores do *Magic Duels* (Magic Duels, 2019). Nele, os jogadores podiam assumir o controle dos personagens e participar de lutas

Commented [EM25]: Acho que não podes mm usar esta palavra, troca a por "encerramento"

Commented [EM26R25]: Caga, afinal podes

Commented [JF27R25]: obrigado n tava a lembrar da palavra

emblemáticas das histórias das coleções lançadas. No lugar dele, foi criado o *Magic Arena*, que não conta mais com este modo história.

Em uma segunda tentativa de explorar mais seus mundos narrativos, foi anunciado o lançamento do Magic: Legends, dessa vez um jogo multijogador de mundo aberto. Entretanto, teve o mesmo destino do *Magic Duels* e em 31 de outubro de 2021, antes mesmo de ser oficialmente lançado, seus servidores foram fechados e os jogadores reembolsados (Bankhurst, 2021).

Por fim, em 2019, uma série de animação ambientada no universo do jogo foi anunciada pela Netflix. Contudo, a série foi cancelada em meio a cortes de gastos por parte da plataforma (Observatório do Cinema, 2024).

Já para os textos de regra, os resultados foram melhores.

Naive Bayes metricas:				
	precision	recall	f1-score	support
R	0.64	0.62	0.63	489
U	0.63	0.54	0.58	488
G	0.75	0.61	0.67	518
W	0.61	0.68	0.64	511
B	0.55	0.69	0.61	531
accuracy			0.63	2537
macro avg	0.64	0.63	0.63	2537
weighted avg	0.64	0.63	0.63	2537
Logistic metricas:				
	precision	recall	f1-score	support
R	0.58	0.64	0.61	489
U	0.65	0.59	0.62	488
G	0.72	0.68	0.70	518
W	0.66	0.71	0.68	511
B	0.64	0.61	0.62	531
accuracy			0.65	2537
macro avg	0.65	0.65	0.65	2537
weighted avg	0.65	0.65	0.65	2537

Com precisão de 63% e 65% com Naive Bayes e Regressão Logística, respetivamente, os modelos possuem uma melhora de aproximadamente 15% na previsão das cores. A cor melhor prevista é o verde com mais de 70% de precisão em ambos os modelos, enquanto o Naive Bayes falha bastante no preto e a Regressão Logística tem pior desempenho na cor vermelha.

Uma precisão de 65% de acerto é considerada bastante aceitável pelo grupo. Este valor nos mostra que a **filosofia das cores** tende a estar presente nos efeitos das cartas. Entretanto,

Commented [EM28]: @João Vitor Figueiredo A acc não era 69% para os dois? Pelo menos é o que está no notebook do isaac

Commented [JF29R28]: ainda n fiz pros dois isso é só text

Commented [EM30R28]: okok

dá a entender que existe uma certa flexibilidade, para que seja possível a criação de novas mecânicas, de modo a que o jogo não fique muito repetitivo.

Embora seja importante para o jogo que as cores tenham objetivos bem definidos, a possibilidade de que novas mecânicas sejam introduzidas, ou que mecânicas pertencentes a uma cor sejam manipuladas para serem encaixadas em outra cor, são importantes para renovar o jogo ao longo das coleções. Isso é uma estratégia que tem como intuito manter a base de jogadores já existente interessada no jogo, induzindo-os a comprar cartas mais recentes ao invés de apenas jogar com aquelas que já possui. De acordo com Zhong e Xu, atualizações que tragam grandes mudanças, quando feitas com menor periodicidade, tal como as coleções de Magic, tendem a aumentar o interesse dos jogadores.



A carta *Chaos Warp*, uma remoção vermelha lançada na coleção "Magic: The Gathering-Commander" (2011)

A carta acima é um exemplo de como outras cores podem mesclar elementos de outras à sua maneira. Devolver cartas para o baralho é uma mecânica associada ao azul, sendo essa a previsão do modelo. Entretanto, a carta é vermelha, por misturar essa mecânica com a aleatoriedade de uma carta nova entrar no campo ao acaso, de acordo com a natureza caótica da cor.

Por fim, foram testados os mesmos modelos utilizando o texto completo da carta combinado em uma única *string*.

Commented [EM31]: Eu acho que isto pode soar um pouco confuso

Commented [DM32R31]: Eu no meu ponto de vista consigo perceber o que o JV quer dizer

Commented [JF33R31]: eu agora reescrevi

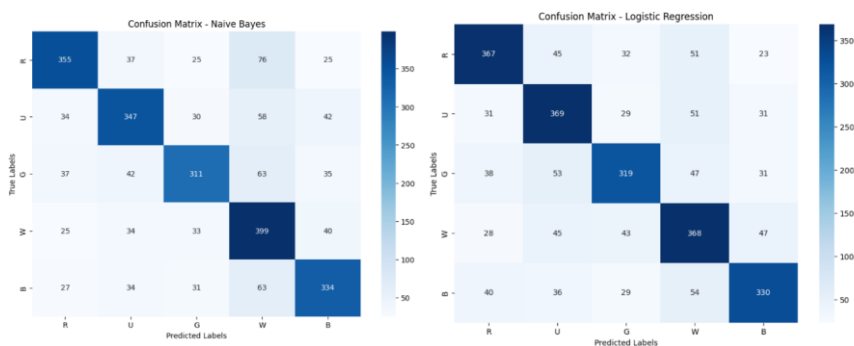
Commented [JF34R31]: tá mais direto ao ponto e embaixo elabora

Model: TF-IDF + Naive Bayes
 Accuracy: 0.69
 Model: TF-IDF + Logistic Regression
 Accuracy: 0.69

Naive Bayes metrics:				
	precision	recall	f1-score	support
R	0.70	0.68	0.69	489
U	0.72	0.64	0.68	488
G	0.74	0.69	0.71	518
W	0.70	0.68	0.69	511
B	0.61	0.75	0.67	531
accuracy			0.69	2537
macro avg	0.70	0.69	0.69	2537
weighted avg	0.69	0.69	0.69	2537

Logistic metrics:				
	precision	recall	f1-score	support
R	0.71	0.67	0.69	489
U	0.71	0.65	0.68	488
G	0.73	0.71	0.72	518
W	0.67	0.72	0.70	511
B	0.64	0.69	0.67	531
accuracy			0.69	2537
macro avg	0.69	0.69	0.69	2537
weighted avg	0.69	0.69	0.69	2537

Este foi o melhor resultado obtido, com 69% de precisão, sendo a pior cor a ser prevista o preto. Os modelos também possuem tendência para falsos positivos no branco, como podemos observar com o suporte da *Confusion Matrix*.



Observando os resultados, a equipe conclui que, embora o lado da criação de histórias tenha um papel fundamental para que o jogo atraia mais jogadores e colecionadores, essas

falhas não foram muito relevantes para a empresa, que soube se adaptar e evitou a perda de jogadores, chegando até a melhorar suas vendas (Schmidt, 2023).

Visto que as cartas em si se mantêm alinhadas com a proposta do jogo, ao mesmo tempo em que existe margem para inovar, e, por consequência, a base de jogadores manteve-se estável, restou à empresa apenas encontrar uma forma de corrigir a falha que foi a criação de histórias e universos.

Seguindo tendências que têm crescido em diversos videogames, como por exemplo *Fortnite*, surgem os *crossovers*. *Crossovers*, neste contexto, são colaborações entre marcas para criação de conteúdos a serem vendidos aos jogadores. No caso do *Fortnite*, existe colaborações com outros videogames ou séries de TV, como *Street-Fighter* e *Marvel*, para criação de cosméticos.



Exemplo de cosméticos criados a partir de *crossover* com a Marvel, retirado da loja virtual *Epic Games*

Seguindo esta tendência, a decisão criativa tomada pela Wizards para coleções futuras de Magic é a de utilizar *crossovers* também.

Commented [EM35]: Atenção, n tenho a certeza de que tenhamos falado já dos Wizards of the coast. Deve-se reforçar que são os detentores do IP



Próximos sets principais para o ano de 2025, retirado de Starcity Games em dezembro de 2024

De acordo com a imagem, podemos observar que a estratégia adotada foi a de intercalar coleções originais com *crossovers*, das 6 coleções principais, ao menos duas serão colaborações (Final Fantasy e Spider-Man). Desta forma, ao aproveitar de universos já prontos, existe maior margem de tempo para equipes criativas trabalharem nas outras coleções. Esta decisão estratégica não só dá margem para trabalhar a falha no lado criativo, mas também atrai novos jogadores, ao inserir conteúdos variados que normalmente não fariam parte do universo de Magic.

Conclusão

Este estudo explorou a aplicação de técnicas de text mining para prever as cores das cartas de Magic: The Gathering com base em seus textos narrativos e mecânicos. A análise demonstrou que elementos textuais fornecem informações valiosas, embora limitadas, para a classificação de cores, refletindo a filosofia das cores do jogo.

A modelagem preditiva alcançou resultados razoáveis, com um modelo combinado atingindo uma precisão de 69%. Os melhores desempenhos foram observados em cartas com características bem definidas, enquanto cartas com atributos mais ambíguos apresentaram desafios para os modelos.

Os insights obtidos indicam que, embora a filosofia das cores esteja presente nos textos das cartas, há espaço para melhorias no alinhamento temático, particularmente em elementos narrativos. Além disso, os resultados sugerem que a flexibilidade na aplicação da **filosofia das cores** contribui para a inovação e renovação do jogo, mantendo o interesse da comunidade.

No geral, este projeto destaca a importância de ferramentas analíticas no estudo de elementos narrativos e mecânicos de jogos de cartas, fornecendo uma base para aprimoramentos nos processos criativos do jogo.

Como trabalho futuro, tendo em conta a necessidade de manter a filosofia das cores no universo Magic, é essencial adaptar as novas abordagens de crossover desenvolvidas pela empresa, preservando a lógica que define a essência das cartas. Para isso, no futuro, será explorado um modelo que aproveite palavras-chaves relevantes dentro das 5 cores do universo Magic, de forma a manter a consistência e coerência com os conceitos fundamentais das cores. Com isso em mente, a abordagem seria utilizar um PMI (Pointwise Mutual Information), para identificar combinações de palavras que capturam a essência das cores clássicas e traduzi-las para novas temáticas, garantindo assim a continuidade da lógica central do jogo. Aproveitar também o PMI para posteriormente se fazer previsões de cartas multi-cor.

Commented [DM36]: O que acham?

Referências bibliográficas

- Zhong, X.; Xu, J. Measuring the effect of game updates on player engagement: A cue from DOTA2. *Entertain. Comput.* **2022**, *43*, 100506
- Common Keywords. MAGIC: THE GATHERING. <https://magic.wizards.com/en/keyword-glossary>. Acesso em dezembro de 2024
- Gendotti, P. (2024). MTG - All Cards. Kaggle.com. <https://www.kaggle.com/datasets/patrickgendotti/mtg-all-cards>. Acesso em dezembro de 2024
- Bankhurst, A. (2021). Magic: Legends is Shutting Down Before Even Leaving Open Beta - IGN. *IGN*. <https://www.ign.com/articles/magic-legends-is-shutting-down-before-even-leaving-open-beta>. Acesso em dezembro de 2024
- Magic Duels. (2019). Magic: The Gathering. <https://magic-support.wizards.com/hc/en-us/articles/360038612991-Magic-Duels>. Acesso em dezembro de 2024
- Observatório do Cinema. (2024). Série de Magic da Netflix é cancelada. *Otempo.com.br*; O TEMPO. <https://www.otempo.com.br/entretenimento/filmes-e-series/2024/9/11/serie-de-magic-da-netflix-e-cancelada>. Acesso em dezembro de 2024
- Sabien, D. (2024). The MTG Color Wheel (& Humanity). *Substack.com*; Homo Sabiens. <https://homosabiens.substack.com/p/the-mtg-color-wheel>. Acesso em dezembro de 2024
- Schmidt, G. (2023). Magic: The Gathering Becomes a Billion-Dollar Brand for Toymaker Hasbro. *The New York Times*. <https://www.nytimes.com/2023/02/16/business/magic-the-gathering-hasbro.html>. Acesso em dezembro de 2024
- Flash, N. (2024). Pop Culture and Gaming: The Rise of Esports and Video Game Crossovers. *Medium*. <https://medium.com/@newsgroup/pop-culture-and-gaming-the-rise-of-esports-and-video-game-crossovers-89a656aed52f>. Acesso em dezembro de 2024