

Estratégia para vencer a última batalha contra seu rival nos jogos da série Pokémon a partir de clusterização

Gabriela Correia Vechini
Imecc UNICAMP
Campinas, Brasil
g172625@dac.unicamp.br

ABSTRACT

Seeking to use machine learning algorithms, we promote a strategy to better choose the set of pokemons used in the final battle of almost every game in the franchise. Taking advantage of the PAM algorithm, we found the centers of the most diverse clusters, with the hypothesis that the more diverse our set in relation to damages and disadvantages due to the types of pokemons, we must have the combination that can counterattack any other set.

KEYWORDS

K-Medoids, K-Means, cluster, pokemon

1 INTRODUÇÃO

Os jogos da série Pokémon foram, e ainda são, populares em todas as faixas etárias e em muitos países. O jogo é do estilo exploração do mapa mundo e rodadas de batalha do estilo "turn-base", ou seja, o jogador e o adversário (normalmente o computador) tem seus turnos bem delimitados para tomar uma ação.

Ao longo dos anos mais e mais pokemons são adicionados aos jogos da série, por vez cada bloco de novos pokemons são chamados de nova "geração", além da expansão dos mapas e havendo novas áreas a serem exploradas. Mas, em suma, o jogo manteve o mesmo estilo desde seu primeiro jogo lançado Pokémon Red e Green em 27 de fevereiro de 1997.

Como parte da fórmula, o jogador, após conseguir seu primeiro pokémon e obter suas pokébolas, pode percorrer o mundo encontrando e capturando novos pokémons para compor sua equipe e enfrentar os líderes dos ginásios e seu rival.

Os ginásios por sua vez possuem um tema, sendo predominantemente de um tipo de pokémon. Todos, com a exceção do seu rival, possuem uma seleção de pokémons do mesmo tipo. Nas batalhas o jogador e o adversário tem um total de 6 pokémons para lutar, a

		TIPO DO POKÉMON DEFENSIVO																	
		Normal	Fire	Water	Grass	Electric	Ice	Fighting	Poison	Ground	Flying	Psychic	Bug	Rock	Ghost	Dragon	Dark	Steel	Fairy
TIPO DO POKÉMON ATACANTE	Normal																		
	Fire																		
	Water																		
	Grass																		
	Electric																		
	Ice																		
	Fighting																		
	Poison																		
	Ground																		
	Flying																		
	Psychic																		
	Bug																		
	Rock																		
	Ghost																		
	Dragon																		
	Dark																		
	Steel																		
	Fairy																		

Figura 1: Tabela de vantagens e desvantagens

partida termina quando todos os pokémons de um dos lados são derrotados.

Possuindo 18 tipos (ou categorias) de pokémons no jogo, o jogador deve fazer escolhas de quais pokémons investir para chegar ao final com mais chances de derrotar o oponente. Sabemos que o tipo de pokémon influencia nas batalhas, pois, alguns tipos tem vantagens e desvantagens contra determinados outros tipos. Podemos observar isto na tabela 1. Essas vantagens e desvantagens variam para cada pokémon, podendo assumir valores entre 0 (seu ataque não causa dano) e 2 (seu ataque é muito efetivo).

Neste trabalho analisaremos uma estratégia para melhor selecionar os pokémons para a última luta. Onde consideramos que o oponente possui uma seleção de pokémons com tipos variados e montando a seleção do jogador com a maior variedade espera-se contra-atacar qualquer pokémon adversário. Assim estando preparados para ter ao menos 1 pokémon que sempre tenha vantagem contra aquele que estamos contra.

2 SOBRE O BANCO DE DADOS

O banco de dados retirado de [1] possui 8 gerações de pokémons e informações sobre seus atributos em batalha, ataque, defesa, pontos de vida, velocidade assim como seu coeficiente de vantagem (de 0 a



Figura 2: PCA

2) contra cada um dos 18 tipos. Considera-se que todos os pokémons estão no nível máximo.

Cada geração introduz em média 112 novos pokémons e as proporções de seus tipo tendem a variar com o tema da nova geração. Alguns pokémons possuem 2 tipos, por exemplo fada e fabatsma ou dragão e fogo, esses pokémons recebem vantagens e desvantagens de ambos dos seus tipos.

Atualmente tem-se o total de 898 pokémons.

3 METODOLOGIA

Neste trabalho vamos agrupar observações semelhantes utilizando a PAN (partition around medoids), criado por Leonard Kaufman e Peter J. Rousseeuw [3].

O PAN é um algoritmo que agrupa observações próximos ao mesmo centro (medoid), visando minimizar uma soma de dissimilaridades de pares em vez de uma soma de distâncias euclidianas quadradas, assim, ele é mais robusto a ruídos e outliers do que k-médias. Em contraste com o algoritmo K-Means, o PAM escolhe uma das observações atuais como centros e, assim, permite maior interpretabilidade dos centros do cluster do que em k-means.

Utilizando 25 variáveis (7 sendo atributos do Pokémon e 18 sua vantagem/desvantagem associada a cada um dos tipos) do banco de dados, foram atualizados os gráficos de cotovelo e silhueta para a decisão da quantidade de clusters a serem utilizados.

No entanto, devido a alta dimensionalidade dos dados, foi usado PCA para reduzir o número de variáveis usadas no modelo. Foi escolhido usar 5 componentes principais pois essas somam cerca de 55,4% da variabilidade explicada, como pode-se observar em Figura 2.

Posteriormente, para visualizar os clusters, utilizou-se a primeira e segunda das componentes principais para se observar em um plano como estão distribuídos os clusters encontrados.

4 RESULTADOS

Observando os gráficos de cotovelo Figura 4 e o gráfico de silhueta Figura 3, a quantidade mais recomendada de clusters seria 3. Pois, é o valor correspondente ao "cotovelo" em Figura 4 e é o melhor valor de Silhouette em Figura 3.

Logo, pode-se visualizar os 3 clusters utilizando as duas primeiras componentes principais na Figura 5.

Para realizarmos computacionalmente o PAN foi feito o uso da função KMedoids da biblioteca Scikit-Learn-Extra, a documentação

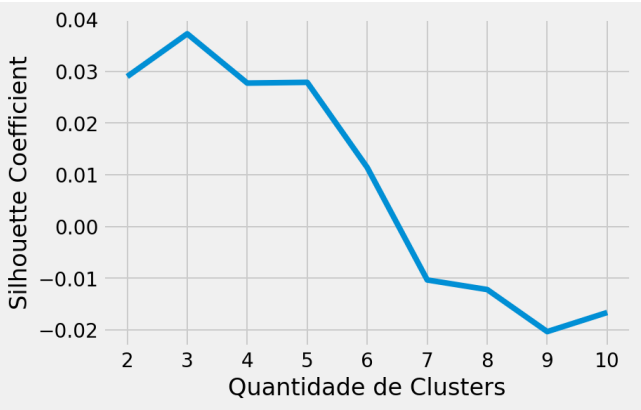


Figura 3: Gráfico de cotovelo

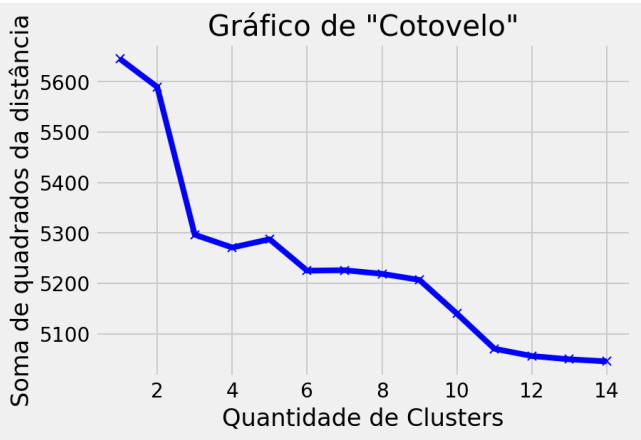


Figura 4: Gráfico de cotovelo

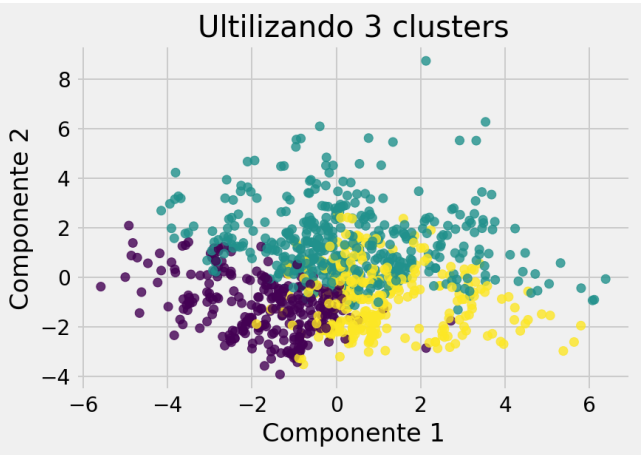


Figura 5: Clusterização com 3 centros

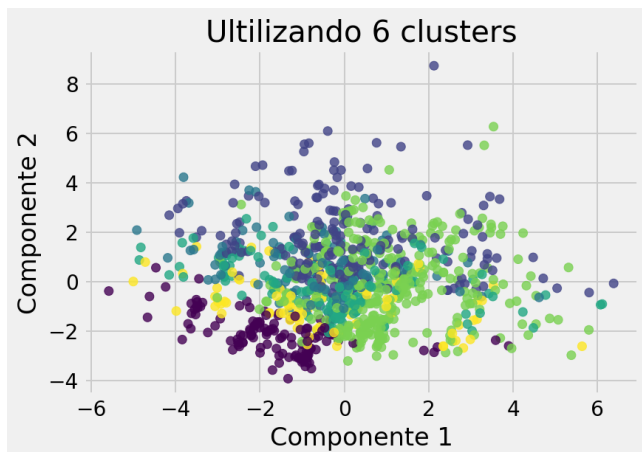


Figura 6: Clusterização com 6 centros

Tabela 1: Time

Id	Nome	Tipo
414	Castform Rainy Form	Water
479	Kricketune	Bug
481	Luxio	Electric
482	Luxray	Electric
702	Eelektross	Electric
900	Minior Meteor Form	Rock Flying

pode ser encontrada em [4]. As demais métricas apresentadas veem da biblioteca Skicit-Learn [5]. Ambas das bibliotecas são para a linguagem Python. Os códigos utilizados podem ser acessados na pagina do GitHub da autora, em [2].

5 DISCUSSÃO

Conforme observado na Figura 5, 3 clusters seriam o "ideal" para agrupar os semelhantes pokémons existentes até agora. Ou seja, poderiam ser separados em 3 grandes grupos com atributos de batalha mais semelhantes.

No entanto, aproveitando-se da interpretabilidade do algoritmo PAM e de que é possível ter até 6 pokémons para a batalha. Seria interessante usar este algoritmo escolhendo como 6 o número de clusters. Observando novamente o Gráfico de cotovelo 4 e de Silhueta 3 percebe-se pouca diferença entre 3 e 6 clusters.

Sendo assim, uma estratégia mais esperta seria usar um pokémon de cada um dos 6 clusters como vemos na Figura 6. Não apenas escolher um pokémon qualquer dentro de cada cluster, mas escolher exatamente aquele que se encontram como centro. Assim, tem-se uma seleção de pokémons mais diversa em referência de centros.

6 CONCLUSÃO

Após discutir a melhor utilização e interpretação dos clusters, temos o time ideal listado na Tabela 1.

REFERÊNCIAS

- [1] Dados retirados de DoltHub: <https://www.dolthub.com/repositories/dolthub/pokemon>. Acesso em: 01/06/2021.
- [2] Código utilizado no repositório da autora: https://github.com/gvechini/ME732_final. Acesso em: 01/06/2021.
- [3] Van der Laan, Mark, Katherine Pollard, and Jennifer Bryan. "A new partitioning around medoids algorithm." *Journal of Statistical Computation and Simulation* 73.8 (2003): 575-584.
- [4] Documentação da função KMedoids em: <https://scikit-learn-extra.readthedocs.io>. Acesso em: 20 abril 2021.
- [5] Documentação da biblioteca Scikit-learn em: <https://scikit-learn.org/stable/>. Acesso em: 20 abril 2021.