Algoritmo de Recomendação de Jogos

Ranniery Dias de Brito

Departamento de Computação – Universidade Federal Rural de Pernambuco (UFRPE) Caixa Postal 52171-900 – Recife – PE – Brasil

ranniery.dias@live.com

**Abstract.** This meta-paper describes the style the Project of the second learning examination of the artificial intelligence discipline, on this paper, I will describe about a game recommendation software based on the user’s previously played games, in the end the application will recommend the game that is most similar based on the attributes and that has the best recommendation rate.

**Resumo.** Este meta-artigo descreve o projeto da segunda avaliação da disciplina de inteligência artificial, neste descreverei sobre um sistema de indicação de jogos baseados nos jogos que o usuário informar, tendo como saída um jogo que melhor corresponde com a entrada do usuário utilizando atributos que mais se assemelha assim como o que tiver maior número em relação ao índice de recomendação.

# Introdução

A Inteligência Artificial (IA) é uma área da computação que tem como objetivo a “automação das atividades que nós associamos com o pensamento humano” (Bellman, 1978), tais como raciocínio, aprendizagem, percepção, entre outras. Portanto, é uma área que, apesar de nova, já está presente em diversas áreas da computação, o que faz com que seu aprendizado seja de grande importância. Sistemas de recomendação por sua vez, são utilizados em vários campos da tecnologia atualmente, desde música, filmes, séries, caminhos a seguir via GPS. Visando fornecer uma alternativa a problemática em relação a escolha de um game, sendo utilizado como base algum ou alguns jogos de preferência do usuário, o sistema então irá recomendá-lo com o exemplar que mais seja compatível com os jogos inseridos, o algoritmo a ser utilizado para a solução será o KNN Ponderado, e até o presente momento não se faz necessário nenhuma modificação para melhoria de resultado

# 1. Objetivo

O objetivo final do projeto é recomendar uma nova opção para o usuário jogar de acordo com a entrada efetuada por este.

# 2. Classes

Como cada jogo é um exemplar individual, o problema em si não possui nenhuma classe, apenas as amostras em questão.

# 3. Características

Para o vetor de características iremos usar os atributos que cada exemplar possui, estes são:

* Categorysingleplayer (Categoria Single Player)
* Categorymultiplayer (Categoria Multi Player)
* Categorycoop (Categoria Cooperativo)
* Categorymmo (Categoria Massive Multiplayer Online)
* Categoryvrsupport (Categoria Suporte a Realidade Virtual)
* Genreisindie (Gênero é indie)
* Genreisaction (Gênero é ação)
* Genreisadventure (Gênero é aventura)
* Genreiscasual (Gênero é casual)
* Genreisstrategy (Gênero é estratégia)
* Genreisrpg (Gênero é Role-Playing Game)
* Genreissimulation (Gênero é simualação)
* Genreissports (Gênero é esportes)
* Genreisracing (Gênero é corrida)
* Genreismassivelymultiplayer (Gênero é multiplayer)
* Recommendationcount (taxa de recomendação)

Além de possuir nome, ID e taxa de recomendação.

# 4. O Dataset

O dataset escolhido foi o *Steam-game-data* (disponível em: <https://data.world/craigkelly/steam-game-data> (acessado em 22/06/2018))

## 5. Algoritmo escolhido

Para a solução proposta, o KNN é um algoritmo de inteligência artificial que visa fazer classificações baseado na distância entre o novo indivíduo e seus vizinhos em um espaço n-dimensional, que pela quantidade de atributos que os exemplares possuem, ou seja, de tamanho 15. Para este caso usaremos o seguinte pseudo-código para o software em questão:

1. **Inicialização:**
2. Recebe o dataset;
3. Receba os jogos do usuário
4. **para** cada novo jogo que o usuário inserir **faça**
5. Calcula a distância da amostra do usuário para as amostras do dataset
6. Determinar os **K** mais próximos
7. O rótulo com maior similaridade e tiver maior índice de recomendação
8. no conjunto dos **K** vizinhos será o escolhido
9. **fim para**
10. **retornar:** jogo escolhido

# 6. Metodologia dos experimentos

A o método utilizado para analisar a similaridade é o índice de similaridade de Hamming para atributos categóricos, que é descrito como:

Além desta foram utilizadas outras 4 distâncias para decisão de que critério para análise de similaridade utilizar, estas foram: Euclidiana, euclidiana ponderada, Manhattan e Manhattan Ponderada.

A linguagem escolhida para tal desenvolvimento foi Java utilizando o paradigma orientado a objetos.

# 7. Resultados

O algoritmo KNN conseguiu fazer a classificação, e após vários testes com 13 exemplares distintos, a distância/atributo categórico que obteve melhores resultados, quanto a precisão de atributos similares, foi a distância de Hamming, onde todos os jogos recomendados continham os mesmos atributos ao qual o jogo de referência.

Quanto ao uso de outras distâncias as recomendações não tiveram resultados tão interessantes, muitas vezes ignorando os atributos de gênero, e levando mais em consideração a taxa de recomendação.

# 8. Análise dos resultados

Considero o resultado satisfatório, tendo em vista que o usuário recebe a recomendação de acordo com os atributos mais semelhantes e a taxa de recomendação mais alta quando o método categórico é a distância de Hamming.

Quanto aos valores que não são tão interessantes quando se é utilizado de outras distâncias, tudo leva a crer que deva ter sido algum equívoco durante o desenvolvimento do código.

# 9. Conclusão

Sistemas de recomendações estão cada vez mais comum na nossa sociedade, e a tendência é que tais ferramentas tenham ainda mais destaque. Ao final do projeto, o algoritmo proposto (KNN) foi desenvolvido, foi utilizado o vetor de características indicado anteriormente. Ainda que recomendação seja um conceito subjetivo, em que não há exatamente uma resposta exata generalizada, o algoritmo conseguiu realizar a recomendação utilizando os critérios propostos. Concluo que outras abordagens possam vir a ser mais interessante como redes neurais, deep learning, aprendizado online, são alguns exemplos, alguma forma de personalizar os dados para cada usuário e aprender com a avaliação do usuário, sendo a utilização de um desses algoritmos nesta temática um possível projeto futuro.