

# Uso de Autoencoders e Clustering para Análise de Jogos de Dota 2

Gustavo F. Ceccon<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista "Júlio de Mesquita Filho"  
Caixa Postal 13506-692 – (19)3526-9000 – Rio Claro – SP – Brasil

`gustavo.ceccon@unesp.br`

**Abstract.** *This article explores the application of autoencoder techniques and clustering methods to identify similar professional matches in Dota 2, one of the main games in the MOBA (Multiplayer Online Battle Arena) genre. Using public match data, the paper proposes an approach based on dimensionality reduction and unsupervised clustering to analyze patterns in hero compositions, economic performance and objectives achieved. The method employs Deep Embedded Clustering (DEC) to simultaneously integrate latent representation learning and clustering, demonstrating the relevance of these techniques for eSports analysis and presenting an overview of the most effective methodologies for clustering similar matches.*

**Resumo.** *Este artigo explora a aplicação de técnicas de autoencoders e métodos de clustering para identificar semelhanças entre partidas profissionais de Dota 2, um dos principais jogos do gênero MOBA (Multiplayer Online Battle Arena). Utilizando dados públicos de partidas, o trabalho propõe uma abordagem baseada em redução de dimensionalidade e agrupamento não supervisionado para analisar padrões em composições de heróis, desempenho econômico e objetivos alcançados. O método emprega Deep Embedded Clustering (DEC) para integrar simultaneamente a aprendizagem de representações latentes e a formação de clusters, demonstrando a relevância dessas técnicas para análise de eSports e apresentando um panorama das metodologias mais eficazes para agrupamento de partidas similares.*

## 1. Introdução

A análise de dados em eSports tem experimentado crescimento devido à riqueza e complexidade dos dados gerados em jogos como Dota 2. O cenário competitivo envolve não apenas jogadores e equipes, mas também analistas, comentaristas e plataformas de estatísticas que buscam prever resultados, identificar padrões estratégicos e fornecer ideias táticas para melhorar o desempenho [Drachen and Schubert 2016]. Empresas bilionárias patrocinam não só os eventos, mas na área de inteligência artificial, da área de tecnologia como Nvidia, desde 2011 quando o jogo estava surgindo. Outras grandes empresas envolvem alimentação, como a Monster Energy e Red Bull, móveis como Secretlab e aposta como 1Bet e GG.Bet, além do patrocínio dos jogadores para a premiação. Milhões de espectadores e premiações milionárias, tem impulsionado a necessidade de métodos analíticos sofisticados para compreender a dinâmica complexa desses jogos [Costa et al. 2023].

Empresas como SAP e a própria desenvolvedora Valve já fazem estatísticas e análises de jogos de jogos e até em tempo real. Dota plus é uma ferramenta que pode ser adquirida no jogo e mostra estatísticas de decisões e vantagens para o jogo. Durante o jogo os analistas conseguem ver a predição de vitória e derrota, ferramenta disponível já implementada dentro do jogo.

Estudos recentes demonstram que a aplicação de aprendizado de máquina em Dota 2 tem sido utilizada para diversas tarefas, incluindo classificação de papéis de jogadores [Eggert et al. 2015], predição de eventos críticos como mortes de heróis [Katona et al. 2019], análise de composições de equipes [Cadman 2024] e detecção de encontros táticos [Schubert et al. 2016]. No entanto, existe uma lacuna na literatura quanto à análise de similaridade entre partidas inteiras, especialmente utilizando técnicas avançadas de autoencoders e clustering profundo, que podem capturar relações não-lineares complexas nos dados de jogos.

O problema central deste trabalho é identificar partidas similares entre si, considerando múltiplas variáveis categóricas e numéricas, com o objetivo de apoiar análises táticas e históricas. Dota 2, como um jogo MOBA complexo, apresenta características únicas que tornam essa análise desafiadora: mais de 120 heróis únicos, milhares de combinações de itens, estratégias emergentes e meta-jogos em constante evolução [?]. A identificação de padrões de similaridade entre partidas pode revelar tendências estratégicas, auxiliar na preparação de equipes e fornecer planos valiosos para análise pós-jogo.

A solução proposta envolve a extração de características relevantes das partidas, pré-processamento dos dados, aplicação de técnicas de redução de dimensionalidade usando autoencoders e agrupamento não supervisionado, seguido de análise supervisionada para validação dos clusters formados. O foco está em jogos profissionais, especialmente partidas de campeonatos como The International, utilizando dados compreendendo o período de 2020 a 2024, capturando assim diferentes versões do jogo e evoluções do meta.

## 2. References

Bibliographic references must be unambiguous and uniform. We recommend giving the author names references in brackets, e.g. [Knuth 1984], [Boulic and Renault 1991], and [Smith and Jones 1999].

The references must be listed using 12 point font size, with 6 points of space before each reference. The first line of each reference should not be indented, while the subsequent should be indented by 0.5 cm.

## Referências

Boulic, R. and Renault, O. (1991). 3d hierarchies for animation. In Magnenat-Thalmann, N. and Thalmann, D., editors, *New Trends in Animation and Visualization*. John Wiley & Sons Ltd.

Cadman, A. (2024). Studying the effects of team composition in role-based competitive video games. Master's thesis, Lancaster University. Department of Computer Science.

- Costa, L. M., Drachen, A., Souza, F. C., and Schön, D. (2023). Artificial intelligence in MOBA games: A multivocal literature mapping. *IEEE Transactions on Games*, 15(3):392–405.
- Drachen, A. and Schubert, M. (2016). Esports analytics through encounter detection. In *MIT Sloan Sports Analytics Conference*.
- Eggert, C., Herrlich, M., Smeddinck, J., and Malaka, R. (2015). Classification of player roles in the team-based multi-player game Dota 2. In *Entertainment Computing–ICEC 2015*, pages 112–125. Springer.
- Katona, A., Spick, R., Hodge, V. J., Demediuk, S., Block, F., Drachen, A., and Walker, J. A. (2019). Time to die: Death prediction in Dota 2 using deep learning. *IEEE Transactions on Games*, 12(3):273–284.
- Knuth, D. E. (1984). *The T<sub>E</sub>X Book*. Addison-Wesley, 15th edition.
- Schubert, M., Drachen, A., and Mahlmann, T. (2016). Esports analytics through encounter detection. In *Proceedings of the MIT Sloan Sports Analytics Conference*.
- Smith, A. and Jones, B. (1999). On the complexity of computing. In Smith-Jones, A. B., editor, *Advances in Computer Science*, pages 555–566. Publishing Press.