

# ***Minority Report*: é possível um modelo computacional prever o comportamento humano?**

**Murilo Augusto Maestro<sup>1</sup>, Helen C. Mattos Senefonte<sup>1</sup>**

<sup>1</sup>Departamento de Computação – Universidade Estadual de Londrina (UEL)  
Caixa Postal 10.011 – CEP 86057-970 – Londrina – PR – Brasil

`murilomaestro@gmail.com, helen@uel.br`

**Abstract.** *In the film *Minority Report*, a narrative is proposed that it would be possible to predict future crimes through a supercomputer. One of the questions the film proposes is: "Is it possible to predict human behavior?". Although we do not have a supercomputer as in the work, we still have the question whether it would be possible to model human behavior, to the point of predicting what the next behavior would be. The work proposed here seeks to exemplify, through performance comparisons, some problems with the film's narrative, demonstrated by the intrinsic characteristics of machine learning models.*

**Resumo.** *No filme *Minority Report*, é proposta uma narrativa em que é possível prever crimes futuros através de um supercomputador. Uma das perguntas que o filme levanta para o espectador é: "Seria possível prever o comportamento humano?". Apesar de não possuímos um supercomputador como na obra, ainda temos a questão de se seria possível modelar o comportamento humano, a ponto de prever o que seria o comportamento próximo. O trabalho aqui proposto visa exemplificar, por comparações de desempenhos, alguns dos problemas com a narrativa do filme, demonstrados pelas características intrínsecas de modelos de machine learning*

## **1. Introdução**

O desenvolvimento de modelos computacionais para o auxílio na predição do comportamento humano pode proporcionar uma ferramenta relevante em muitos contextos, tais quais, prever o fluxo de turistas em uma cidade [11], saber quais produtos oferecer a um indivíduo [13], identificar crimes em potencial [1], entre outros. Essas possibilidades de aplicação possuem características problemáticas que podem ser auxiliados por modelos computacionais seja oferecendo intuições sobre os dados processados, levantando estimativas, dentre outras formas.

Diversos modelos são concebidos, desenvolvidos e estudados a fim de oferecer previsões sobre dados de entrada. As abordagens propostas na literatura utilizam diferentes técnicas conjuntas para extrapolar os dados necessários objetivando as previsões de comportamentos. Com tantas opções de modelos é comum levantar a questão, qual o melhor modelo para um contexto arbitrário? E ainda outra pergunta, seria possível comparar modelos sendo que foram concebidos para resolver diferentes problemas? E com isso seria possível responder à pergunta: é possível modelar o comportamento humano?

O presente projeto é um estudo de três (3) outros trabalhos e seus respectivos modelos de predição de comportamento, e a adaptação de seus algoritmos em contextos

diferentes. Serão levantados dados comparativos do desempenho individual, a fim de avaliar seus desempenhos e características e motivos de suas respectivas qualidades em cada contexto.

## **2. Fundamentação Teórico-Metodológica e Estado da Arte**

Nessa seção, são apresentados os principais modelos usados na literatura em problemas diversos de predição.

### **2.1. *Self Organizing Maps with Fuzzy C***

Em [11] é utilizado uma base dados sobre o deslocamento de turistas em uma viagem. Descritores de movimento são construídos a partir das transições de categorias de locais que o individuo visita, e através desses conjuntos é criado uma trajetória de lugares pelo que tempo que o turista esteve.

A técnica de *Self Organizing Maps* (SOMs), ou Mapas Auto Organizáveis, ou ainda Mapas de Kohonen [6], é implementada em conjunto com outra técnica de classificação proposta em [3]. Esta concatenação forma um método que permite o uso da organização espacial gerada pelo SOM (Self Organizing Map) de forma que seja possível definir o grau de pertencimento de um dado turista tem a diferentes perfis.

### **2.2. *Factorization Machines***

O comportamento modelado em [13] reflete a interação de um comprador com um site de vendas de varejo. O trabalho adapta o modelo apresentado em [10], chamado *Factorization Machine*, para o contexto proposto, fazendo ajustes onde necessário. Este modelo faz a combinação das vantagens *Support Vector Machines* (SVMs), com formas de fatorização de dados, permitindo o processamento de dados com muitas dimensões em tempo linear. Outra vantagem do uso deste modelo é a capacidade de manter acurácia com dados esparsos.

### **2.3. *Recursive Clustering with Linear Regression***

A fim de auxiliar na previsão de estudantes que podem ter dificuldades nos estudos acadêmicos, o trabalho [2] propõe um modelo que pode ajudar a prever quais alunos podem precisar de apoio nos estudos. A técnica de *Recursive clustering*, similar à [6], é aplicada para se obter três (3) grupos de estudantes, abaixo da média, dentro da média e acima da média, utilizados para selecionar candidatos para o apoio acadêmico, e terem seus próximos resultados preditos por uma algoritmo de regressão linear.

### **2.4. Trabalhos Correlatos**

Os trabalhos correlatos citados nessa seção, buscam propor modelos capazes de prever decisões feitas por humanos, tendo como base decisões prévias, ou até sem qualquer outros dado de comportamento prévio. Esses modelos desenvolvidos usam diversas técnicas em conjunto e/ou em partes separadas de seus processos para produzir resultados.

Em [9] busca e analisa diferentes modelos utilizados para predição de desempenho de estudantes em diversos níveis de ensino. Também explica e levanta dados sobre o desempenho de diferentes modelos de predição para o desempenho acadêmico. Outro trabalho no tema de ensino e auxílio a alunos é [4], que propõe uma estrutura de *machine*

*learning* que visa sugerir diferentes metodologias de estudo para estudantes tendo como base desempenhos prévios.

Buscando trabalhar em um tema mais próximo de assuntos de psicologia, os trabalhos [5], [7] e [12] avaliam o desempenho de técnicas de modelagem computacionais comparadas com técnicas estatísticas e outras aplicadas de forma classificatória para prever suicidas em potencial.

Pesquisar temas relacionados ao comportamento humano e sistemas de recomendação é possível encontrar os trabalhos [8], [13] e [11]. Os dois primeiros propõem diferentes modelos para prever os comportamentos de compradores em potencial em sites de varejo, enquanto o último expõe uma estrutura de *machine learning* que visa conseguir prever o caminho de um turista através de uma cidade que visita.

Dentre os vários trabalhos comentados no projeto, os selecionados para o estudo principal são:

- PredicTour: Predicting mobility patterns of tourists based on social media user's profiles. [11]
- The application of factorization machines in user behavior prediction. [13]
- Recursive clustering technique for students' performance evaluation in programming courses. [2]

### **3. Objetivos**

Comparar diferentes modelos de *machine learning* que buscam prever comportamentos e/ou resultados de comportamento humano. Através de contextos diferentes exemplificar as qualidades e peculiaridades de cada modelo, denotando as adaptações necessárias, assim como limitações existentes pelos modelos.

### **4. Procedimentos metodológicos/Métodos e técnicas**

Primeiramente, o estudo dos modelos mencionado nas subseções 2.1, 2.2 e 2.3, é essencial para a definição de quais algoritmos serão usados, seus parâmetros e pontos de atenção, esboçando possíveis desafios de implementação. Será realizada então a seleção de dados (*datasets*) interessantes para o estudo, considerando possíveis fatores limitantes levantados na primeira etapa.

Serão feitas as implementações de cada modelo, adaptadas aos contextos dos dados, e seus desempenhos anotados. E por fim, é esperado através da comparação dos desempenhos e comportamentos de cada modelo em cada contexto explorado, demonstrar as diferentes qualidades de cada modelo. A seguir uma lista de atividades é proposta e organizada cronologicamente na Tabela 1 (Cronograma de Execução),

Atividades:

1. Revisão Bibliográfica;
2. Estudo de técnicas e modelos computacionais para previsão de comportamento humano;
3. Entrega de Projeto de TCC;
4. Redação de estudos realizados;
5. Implementação de métodos em contexto selecionados;

6. Implementação de métodos alternativos;
7. Comparação de resultados obtidos;
8. Redação de TCC final;

**Tabela 1. Cronograma de Execução**

	jul	ago	set	out	nov	dez
Atividade 1	X	X	X			
Atividade 2		X	X			
Atividade 3			X			
Atividade 4			X	X	X	
Atividade 5				X		
Atividade 6					X	
Atividade 7						X
Atividade 8						X

### 5. Contribuições e/ou Resultados esperados

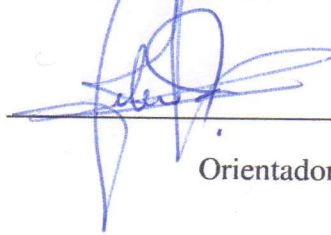
Com este trabalho, espera-se desenvolver um estudo comparativo dos modelos, propostos na literatura, para previsão de comportamentos. Também espera-se que através do trabalho seja possível exemplificar uma metodologia para comparação de diferentes modelos de *machine learning* focados na previsão de comportamento humano.

### 6. Espaço para assinaturas



Aluno

Londrina, 12 de setembro de 2022.



Orientador

## Referências

- [1] Zaheer Abbass, Zain Ali, Mubashir Ali, Bilal Akbar, and Ahsan Saleem. A framework to predict social crime through twitter tweets by using machine learning. In *2020 IEEE 14th International Conference on Semantic Computing (ICSC)*, pages 363–368, 2020. doi:10.1109/ICSC.2020.00073.
- [2] V. K. Anand, S. K. Abdul Rahiman, E. Ben George, and A. S. Huda. Recursive clustering technique for students’ performance evaluation in programming courses. In *2018 Majan International Conference (MIC)*, pages 1–5, 2018. doi:10.1109/MINTC.2018.8363153.
- [3] James C. Bezdek, Robert Ehrlich, and William Full. Fcm: The fuzzy c-means clustering algorithm. *Computers Geosciences*, 10(2):191–203, 1984. doi:/10.1016/0098-3004(84)90020-7.
- [4] Andi Besse Firdausiah Mansur, Norazah Yusof, and Ahmad Hoirul Basori. Personalized learning model based on deep learning algorithm for student behaviour analytic. *Procedia Computer Science*, 163:125–133, 2019. 16th Learning and Technology Conference 2019 Artificial Intelligence and Machine Learning: Embedding the Intelligence. doi:10.1016/j.procs.2019.12.094.
- [5] Leandro Nicolás Grendas, Luciana Chiapella, Demian Emanuel Rodante, and Federico Manuel Daray. Comparison of traditional model-based statistical methods with machine learning for the prediction of suicide behaviour. *Journal of Psychiatric Research*, 145:85–91, 2022. doi:10.1016/j.jpsychires.2021.11.029.
- [6] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990. doi:10.1109/5.58325.
- [7] Catherine M McHugh and Matthew M Large. Can machine-learning methods really help predict suicide? *Current opinion in psychiatry*, 33(4):369–374, 2020. doi:10.1097/YCO.0000000000000609.
- [8] Adebola Orogun and Bukola Onyekwelu. Predicting consumer behaviour in digital market: a machine learning approach. 2019. doi:10.15680/IJIRSET.2019.0808006.
- [9] Juan L. Rastrollo-Guerrero, Juan A. Gómez-Pulido, and Arturo Durán-Domínguez. Analyzing and predicting students’ performance by means of machine learning: A review. *Applied Sciences*, 10(3), 2020. doi:10.3390/app10031042.
- [10] Steffen Rendle. Factorization machines. In *2010 IEEE International conference on data mining*, pages 995–1000. IEEE, 2010.
- [11] Helen C. Mattos Senefonte, Myriam Regattieri Delgado, Ricardo Lüders, and Thiago H. Silva. Predictour: Predicting mobility patterns of tourists based on social media user’s profiles. *IEEE Access*, 10:9257–9270, 2022. doi:10.1109/ACCESS.2022.3143503.
- [12] Kasper van Mens, CWM de Schepper, Ben Wijnen, Saskia J Koldijk, Hugo Schnack, Peter de Loeff, Joran Lokkerbol, Karen Wetherall, Seonaid Cleare, Rory C O’Connor, and Derek de Beurs. Predicting future suicidal behaviour in young adults, with different machine learning techniques: A population-based longitudinal study. *Journal*

*of Affective Disorders*, 271:169–177, 2020. doi:10.1016/j.jad.2020.03.081.

- [13] Yuqi Wang, Wenqian Shang, and Zhenzhong Li. The application of factorization machines in user behavior prediction. In *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, pages 1–4, 2016. doi:10.1109/ICIS.2016.7550927.