



www.aileaders-project.eu

SIMULAÇÃO - Desafios em Recursos Humanos



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

SIMULAÇÃO

- Desafios éticos em Recursos Humanos

01



Resumo 3

02



Introdução 4

03



Apresentação das ferramentas 6

04



Execução da simulação 7

05



Conclusões 8

06



Referências 9



Tipo de REA

Demonstração/Simulação utilizando o Google Colab (RandomForest vs FairGBM Classifier)

Objetivo/Finalidade

Proporcionar aos estudantes uma exploração prática e crítica de como a tomada de decisões algorítmica no recrutamento pode reproduzir desigualdades estruturais, comparando os resultados de métricas de equidade. Esta simulação incentiva a reflexão sobre as dimensões éticas da IA nos Recursos Humanos e promove o desenvolvimento de modelos preditivos mais justos e inclusivos.

Resultados de aprendizagem esperados

Ao final da simulação, os alunos serão capazes de:

- 01** Detetar e interpretar o viés algorítmico na contratação impulsionada pela IA;
- 02** Refletir sobre as implicações éticas dos sistemas de recrutamento automatizados
- 03** Utilizar métricas de equidade para avaliar os resultados do modelo;

Palavras-chave

- Aprendizagem automática
- Classificação
- Recursos humanos
- Preconceitos
- Equidade

Sugerido Abordagem metodológica

Aprendizagem baseada em problemas

• 02 Introdução



Glossário

- 01** **Responsabilidade** significa que a organização é responsável pelas suas decisões de contratação e pelos seus resultados, garantindo que existam consequências para práticas injustas.
- 02** **Transparência** envolve tornar o processo de recrutamento e os critérios de seleção claros e compreensíveis para todos os candidatos.
- 03** **A equidade** garante que todos os candidatos sejam avaliados de forma imparcial com base nas suas qualificações para a função, sem qualquer preconceito ou discriminação.
- 04** **O preconceito social** na contratação é a tendência injusta de favorecer ou rejeitar candidatos com base em suposições preconceituosas sobre o seu grupo social (como idade, género ou raça), em vez de suas habilidades e qualificações reais.

Espera-se cada vez mais que os processos de recrutamento sejam apoiados por tecnologias de Inteligência Artificial (IA), particularmente nas fases iniciais da contratação - tais como a triagem de currículos, a classificação de candidatos e a pré-seleção de candidatos para entrevistas. Embora estes sistemas ofereçam o potencial para uma maior eficiência e escalabilidade, também levantam importantes questões éticas, particularmente no que diz respeito à equidade e ao preconceito (Horodyski, 2023).



Muito antes da adoção da IA, pesquisas sobre práticas tradicionais de recrutamento e seleção já haviam documentado padrões persistentes de discriminação (Breugh, 2013; Hebl et al., 2020). Isso inclui preconceitos relacionados à raça, etnia e status de minoria (por exemplo, Allen & Vardaman, 2017; Hiemstra et al., 2013; Veit & Thijsen, 2019; Zschirnt & Ruedin, 2016); género (por exemplo, Ellemers, 2018); classe social (por exemplo, Henderson, 2018); idade (por exemplo, Czopp et al., 2015; Zaniboni et al., 2019); e formação académica (por exemplo, Daly et al.,

2000).

Espera-se cada vez mais que os processos de recrutamento sejam apoiados por tecnologias de Inteligência Artificial (IA), particularmente nas fases iniciais da contratação — tais como triagem de currículos, classificação de candidatos e pré-seleção de candidatos para entrevistas, o que pode levantar importantes questões éticas, particularmente no que diz respeito à equidade e ao preconceito (Horodyski, 2023).

Muito antes da adoção da IA, pesquisas sobre práticas tradicionais de recrutamento e seleção já haviam documentado padrões persistentes de discriminação (Hebl et al., 2020). Isso inclui preconceitos relacionados a raça, idade, etnia e status de minoria (por exemplo, Allen & Vardaman, 2017; Veit & Thijsen, 2019; Zaniboni et al., 2019).



A preocupação agora é que as tecnologias de IA, em vez de mitigar esses preconceitos, possam consolidá-los ou até mesmo ampliá-los. Estudos recentes sugerem que a tomada de decisões algorítmicas em recursos humanos pode replicar as desigualdades históricas incorporadas nos dados utilizados para treinar esses sistemas (por exemplo, Rigotti & Fosh-Villaronga, 2024; Seppälä & Małecka, 2024). Como resultado, o uso da IA no recrutamento levanta questões críticas sobre responsabilidade, transparência e justiça nas práticas de contratação.

Neste contexto, o presente exercício de simulação foi concebido para envolver os estudantes nos desafios éticos do recrutamento e da seleção de candidatos impulsionados pela IA. O objetivo é proporcionar uma experiência prática na identificação de potenciais preconceitos nos

algoritmos de contratação e na exploração de estratégias conscientes da equidade para apoiar uma tomada de decisão mais inclusiva e equitativa.

Através da utilização de métricas de equidade, os estudantes analisarão se os modelos de aprendizagem automática utilizados para prever decisões de pré-seleção refletem — ou reproduzem — padrões históricos de discriminação, com especial atenção à etnia, gênero e contexto socioeconômico.

Ao enfrentar esses desafios em um ambiente controlado e simulado, os estudantes desenvolverão tanto a competência técnica quanto a sensibilidade ética essenciais para o projeto e a implantação responsáveis da IA na gestão de recursos humanos.



• 03 Apresentação das ferramentas



Esta simulação aborda uma tarefa de classificação binária no recrutamento baseado em IA: prever se um candidato deve ser pré-selecionado para um cargo de gestão de nível médio com base no seu perfil.

O conjunto de dados utilizado é sintético, mas modelado com base em dados reais de recrutamento, disponíveis publicamente através da plataforma Kaggle. Os atributos dos candidatos incluem género, raça/etnia, nível de escolaridade, anos de experiência, expectativas salariais e fonte de recrutamento.

Embora o conjunto de dados seja artificial, ele simula de forma realista decisões e padrões comuns de contratação, tornando-o ideal para estudar a equidade na aprendizagem máquina. É particularmente preocupante a forma como variáveis sensíveis, como raça/etnia, podem influenciar as previsões do modelo. Estudos

demonstraram que os processos de recrutamento podem sistematicamente prejudicar candidatos de determinadas origens étnicas, mesmo quando as qualificações são iguais (Zschirnt & Ruedin, 2016; Veit & Thijsen, 2019; Hiemstra et al., 2013).

Esta simulação centra-se numa **questão principal**:
Os algoritmos de ML “recomendam” de forma desproporcional a rejeição de candidaturas a empregos apresentadas por negros e mulheres?

Ao investigar esta questão, os estudantes avaliam criticamente se os modelos de IA utilizados na seleção de candidatos apresentam preconceitos não intencionais — e como esses padrões se relacionam com as conclusões empíricas da literatura académica.

04 Execução da simulação



01

Acesse o Caderno de Simulação

Aceda a <https://tinyurl.com/k63793wp>

02

Execute o código

- Para executar todas as células, clique no separador «Runtime» e selecione «Run All». Aguarde algum tempo (1 minuto).
- Certifique-se de que as saídas carreguem corretamente e que todos os modelos sejam treinados com sucesso.

03

Explore o conjunto de dados

- Na ferramenta What If, clique em «Performance and Metrics» (Desempenho e métricas) e, em seguida, selecione o recurso «Over 50k» (Mais de 50 mil) em «Ground Truth Feature» (Recurso de verdade fundamental) e escolha «GenderCode» (Código de género) para «Slice by» (Fatiar por).
- Analise o conjunto de dados e inspecione as variáveis principais. Preste atenção especial às características demográficas, como raça/etnia e género, e sua possível correlação com os resultados da pré-seleção.

04

Analise a equidade

Aplique métricas de equidade para analisar as diferenças nas taxas de precisão de rejeição entre grupos demográficos, nomeadamente raça/etnia.



Datapoint editor

Performance & Fairness

Features

600 datapoints loaded

Configure

Ground Truth Feature

Over-50K

Cost Ratio (FP/FN)

1

Slice by

RaceDesc

Slice by (secondary)

<none>

WHAT IS GROUND TRUTH?

The feature that your model is trying to predict. [More](#).

WHAT IS COST RATIO?

The cost of false positives relative to false negatives. Required for optimization. [More](#).

WHAT DOES SLICING DO?

Shows the model's performance on datapoints grouped by each value of the selected feature.

Fairness

Apply an optimization strategy

Select a strategy to automatically set classification thresholds, based on the set cost ratio and data slices. Manually altering thresholds or changing cost ratio will revert the strategy to 'custom thresholds'.

☒ Custom thresholds

☐ Single threshold

☐ Demographic parity

☐ Equal opportunity

☐ Equal accuracy

☐ Group thresholds

Custom thresholds for 5 values of RaceDesc

Sort by

Count

Feature Value	Count	Model	Threshold	False Positives (%)	False Negatives (%)	Accuracy (%)	F1
▶ Hispanic	125	1		0.5	0.0	24.0	76.0
	2			0.5	3.2	20.8	76.0
▶ Other	124	1		0.5	0.0	28.2	71.8
	2			0.5	3.2	25.0	71.8
▶ Black	122	1		0.5	0.0	23.8	76.2
	2			0.5	4.9	18.9	76.2
▶ Asian	117	1		0.5	0.0	24.8	75.2
	2			0.5	5.1	19.7	75.2
▶ White	112	1		0.5	0.9	28.6	70.5
	2			0.5	8.0	21.4	70.5

05 Compare os resultados

Compare os resultados entre os grupos, particularmente no que diz respeito à etnia. Determine se candidatos igualmente qualificados de diferentes origens étnicas recebem tratamento desigual. Reflita sobre como isso reflete padrões conhecidos de discriminação na contratação.

06 Reflita sobre as implicações éticas

Discuta se e como o modelo reflete ou reforça o preconceito social. Avalie a eficácia das intervenções conscientes da equidade. Considere implicações éticas mais amplas para o uso da IA no recrutamento e gestão de talentos.

• 05 Conclusão



Esta simulação ilustra como os modelos de aprendizagem automática, se não forem verificados, podem incorporar e reproduzir preconceitos históricos nos processos automatizados de recrutamento. Embora o conjunto de dados utilizado seja sintético e o modelo intencionalmente simplificado, as conclusões refletem preocupações do mundo real em torno da equidade na contratação, particularmente no que diz respeito às disparidades de género e étnicas.

Os resultados revelam inconsistências significativas na precisão preditiva entre grupos demográficos. Especificamente, a precisão das previsões de rejeição de recrutamento é notavelmente maior para mulheres negras. Neste

contexto, maior precisão corresponde a uma taxa mais baixa de falsos negativos, o que significa que as candidaturas de mulheres negras são mais frequentemente e corretamente classificadas como rejeições.

Resumo dos resultados



Precisão na previsão de decisões de rejeição de recrutamento — ou seja, verdadeiros «negativos» (%)

Hispânico

**74
MULHER**

**82
HOMEM**

Negro

**88
MULHERES**

**68
MASC**

Branco

**76
MULHER**

**69%
HOMENS**

Isso sugere que o modelo “recomenda” de forma desproporcional a rejeição de candidatos desse grupo, quando comparado com os de outros grupos, destacando um viés preocupante no processo de tomada de decisão do algoritmo.

Essas descobertas reacendem discussões críticas sobre pontos cegos na contratação algorítmica, particularmente as maneiras pelas quais as desigualdades estruturais são codificadas nos dados de treino e, subsequentemente, aprendidas pelos sistemas de IA. A observação de que a precisão preditiva varia entre grupos marginalizados — e mesmo dentro de categorias protegidas — ressalta a complexidade de avaliar a justiça: maior precisão não indica necessariamente resultados mais equitativos.

Ao aplicar métricas de justiça e examinar os resultados ao nível do grupo, os estudantes são convidados a envolver-se criticamente com as dimensões sociotécnicas do preconceito algorítmico. Este exercício não só proporciona experiência prática com técnicas de detecção e mitigação de preconceitos, como também enfatiza a responsabilidade ética dos programadores e analistas na concepção de sistemas de IA justos e responsáveis. No contexto da gestão de recursos humanos, a justiça não é apenas uma questão legal ou de reputação — é uma obrigação ética fundamental.

• 06 Referências



- Allen, D. G., & Vardaman, J. M. (2017). Recrutamento e retenção em diferentes culturas. *Revisão Anual de Psicologia Organizacional e Comportamento Organizacional*, 4(1), 153–181.
- Breugh, J. A. (2013). Recrutamento de funcionários. *Revisão Anual de Psicologia*, 64(1), 389–416.
- Czopp, A. M., Kay, A. C., & Cheryan, S. (2015). Os estereótipos positivos são generalizados e poderosos. *Perspectivas sobre a ciência psicológica*, 10(4), 451-463.
- Daly, M. C., Büchel, F., & Duncan, G. J. (2000). Prêmios e penalizações para o excesso e o déficit de educação: evidências dos Estados Unidos e da Alemanha. *Revisão da Economia da Educação*, 19(2), 169-178.
- Ellemers, N. (2018). Estereótipos de gênero. *Revisão Anual de Psicologia*, 69(1), 275-298.
- HeModern Discrimination in Organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, M., Cheng, S. K., & Ng, L. C. (2020)., 7(1), 257–282.
- Henderson, D. (2018). Os efeitos da classe social nas percepções sobre a adequação dos candidatos a emprego. *Anais da Academia de Gestão*, 2018(1), 13748.
- Hiemstra, A. M. F., Derous, E., Serlie, A. W., & Born, M. P. (2013). Efeitos da etnia no conteúdo do currículo dos graduados. *Psicologia Aplicada*, 62(3), 427–453.
- Horodyski, P. (2023). Percepção dos recrutadores sobre as ferramentas baseadas em inteligência artificial (IA) no recrutamento. *Relatórios sobre computadores no comportamento humano*, 10, 100298.
- Rigotti, C., & Fosch-Villaronga, E. (2024). Equidade, IA e recrutamento. *Revisão de Direito e Segurança Informática*, 53, 105966.
- Seppälä, P., & Małecka, M. (2024). IA e decisões discriminatórias no recrutamento: desafiando os pressupostos fundamentais. *Big Data & Society*, 11(1), 20539517241235872
- Veit, S., & Thijsen, L. (2019). Quase idênticos, mas ainda assim tratados de forma diferente: discriminação na contratação contra minorias nascidas no estrangeiro e no país. *Journal of Ethnic and Migration Studies*, 1–20.
- Zaniboni, S., Kmicinska, M., Truxillo, D. M., Kahn, K., Paladino, M. P., & Fraccaroli, F. (2019). Ainda me contrataria se eu tivesse mais de 50 anos? Os efeitos dos estereótipos implícitos e explícitos relacionados à idade na avaliação de currículos. *European Journal of Work and Organizational Psychology*, 28(4), 453-467.
- Zschirnt, E., & Ruedin, D. (2016). Discriminação étnica nas decisões de contratação: uma meta-análise de testes de correspondência 1990–2015. *Jornal de Estudos Étnicos e Migratórios*, 42(7), 1115–1134.



leaders

Acompanhe a nossa jornada



www.aileaders-project.eu



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.