

Universidade Aberta

Ética e Práticas de Engenharia

Caso 2 – Conflito entre ética profissional e pressão comercial

Nº Aluno	Nome	Cor
2404582	Pedro Aguilar	
2404398	Sara Coelho	
2405281	André Santos	
2404716	Elen Oliveira	
2404386	Pedro Almeida	
2404498	Diogo Barbosa	
2404699	Marta Baptista	

Pergunta 1 – Descrição do problema e atores envolvidos

Maria é uma engenheira de machine learning numa startup que desenvolveu um algoritmo de triagem de currículos. Durante os testes finais, a Maria e um colega descobrem que o algoritmo está a dar pontuações mais baixas a candidatos de certos grupos, porque foi treinado com dados antigos que já tinham preconceitos, baseados em práticas históricas de contratação injustas.

A liderança da startup está ansiosa para lançar o produto devido à pressão dos investidores e da concorrência. Maria enfrenta o dilema de ignorar o viés (permitindo o lançamento no prazo) ou reportá-lo e exigir a sua correção, o que atrasará o produto, consumirá recursos e poderá colocar em risco a viabilidade financeira da startup e a posição da própria equipa.

Trata-se de um típico “vácuo de política” (James Moor): a tecnologia avançou mais rápido do que as políticas éticas da empresa. Maria tem o conhecimento técnico e a responsabilidade, pois os computadores são “logicamente maleáveis” e cabe aos profissionais moldá-los corretamente.

Os atores principais são:

- Maria e colega (equipa técnica);
- Liderança da startup e investidores;
- Candidatos de grupos minoritários (lesados diretos);
- Empresas clientes, que poderão adotar o sistema sem saber do enviesamento, amplificando o impacto social;
- A sociedade em geral também é afetada, pois a perpetuação da discriminação digital tem impacto estrutural sobre o mercado de trabalho.

Pergunta 2 – Prejuízo, justiça, preconceito e razoabilidade

O **prejuízo** é em primeiro lugar social, pelo reforço de desigualdades, limitação da diversidade nas empresas e exclusão dos candidatos afetados; e em segundo profissional: risco de boicotes, processos judiciais (ex.: RGPD, leis antidiscriminação) e perda irreversível de confiança dos clientes.

Estamos perante aquilo que Huff chama de “Poder Não Intencional” (Huff, Livro 1, Cap. 4, p. 49): o viés não é malicioso nem intencional, mas o poder de excluir minorias é real e prejudicial, o que acaba por constituir preconceito. Ignorar deliberadamente este dano viola o teste do “Ponto de Vista Ético” de Bynum e o princípio de Fairness. Gotterbarn reforça: culpar os dados históricos ou dizer que é “erro de computador” é antiético.

Não existe **justiça**: candidatos competentes de certos grupos nunca chegarão a entrevista, mesmo tendo competência para tal.

O **preconceito** é herdado dos padrões históricos nos dados de treino. Após a descoberta consciente do viés, mantê-lo não é razoável – o risco jurídico, ético e reputacional supera largamente qualquer benefício financeiro do lançamento apressado.

Um lançamento precipitado pode parecer razoável a curto prazo, mas é insustentável a médio prazo e nem existe **razoabilidade** em comprometer a legitimidade dos processos de recrutamento dos clientes e expor a startup a riscos legais e comerciais muito maiores, pois contraria o dever fundamental da engenharia de promover justiça e bem-estar humano (Ética para Engenheiros, p. 57-60).

Pergunta 3 – Conclusões provisórias e questões éticas envolvidas

Maria deve aplicar a não maleficência (evitar causar dano mesmo com custos), justiça, equidade e responsabilidade profissional positiva, servindo o bem comum acima do lucro imediato.

Por isso, mesmo correndo riscos, Maria deve reportar formalmente o erro. Ignorar torná-la cúmplice com um proceder discriminatório, perderia a confiança das chefias e iria deixá-la numa posição frágil em todos os sentidos, inclusive legal. Por outro lado, a empresa veria a sua imagem e reputação afetadas com danos possivelmente irreversíveis.

Ao lançar sem corrigir, a equipa torna-se coautora da discriminação (responsabilidade coletiva). Há também obrigação profissional de validar sistemas que afetam diretamente pessoas.

Impõe-se a “Responsabilidade Positiva”. Lançar a aplicação violaria o Imperativo 1.4 do Código da ACM (“a discriminação não será tolerada”). O caso também ilustra o conflito entre ética deontológica (agir segundo o dever moral) e consequencialista (avaliar resultados). Questão central: a lealdade à empresa pode sobrepor-se ao dever de justiça social?

Pergunta 4 – Como o grupo resolveria a situação

O primeiro passo seria documentar de forma sistemática e rigorosa o viés identificado, incluindo métricas quantitativas, exemplos reais de casos afetados, comparação entre grupos demográficos e impacto estatístico. Este relatório técnico serviria como base para justificar decisões éticas e técnicas, reduzindo espaço para interpretações subjetivas ou resistência da liderança.

O grupo apresentaria o relatório à liderança propondo o adiamento do lançamento automático. A proposta incluiria um plano estruturado com etapas de reequilíbrio dos dados, técnicas de fairness-aware machine learning, criação de novos testes de validação e critérios de aprovação antes da automatização. O objetivo seria garantir que o sistema só seria lançado quando respeitasse princípios de equidade e não maleficência.

Como compromisso temporário, caso existisse pressão económica extrema, poderiam ser aplicados patches corretivos iniciais. Estes seriam acompanhados de métricas de controlo contínuo para avaliar o comportamento do sistema em tempo real, garantindo que mesmo uma implementação parcial não reproduzisse discriminação. A equipa também recomendaria reuniões semanais para revisão ética e técnica do progresso.

Antes do lançamento, seria fundamental ajustar os dados de treino ou, se necessário, alterar o próprio modelo. Isto inclui remover atributos correlacionados com discriminação, aplicar técnicas de mitigação de viés, revendo pesos de variáveis e realizando testes A/B comparando a versão original com a corrigida para garantir melhoria significativa.

A equipa sugeriria implementar um lançamento piloto restrito, direcionado a clientes selecionados, permitindo recolher feedback real sem comprometer a reputação global da empresa. Além disso, seria desenvolvido um módulo de transparência que explicasse aos candidatos como as pontuações foram geradas, garantindo accountability. Seriam também definidas métricas obrigatórias de fairness, aplicadas regularmente como condição de manutenção do sistema.

Outra medida essencial seria comunicar de forma completamente transparente com investidores. O viés não deve ser tratado como simples atraso técnico, mas como defeito crítico com implicações legais e éticas. Esta abordagem reforçaria a responsabilidade social da empresa, evitando riscos futuros. A equipa recomendaria preparar um documento formal que demonstrasse os riscos jurídicos associados a um lançamento prematuro.

Por fim, o grupo defenderia auditorias externas e independentes ao algoritmo e ao processo de recrutamento automatizado, garantindo que a startup se ajusta às melhores práticas internacionais. Também seria promovida formação interna sobre ética em IA, para que a equipa compreenda plenamente os riscos e responsabilidades. O lançamento só aconteceria quando o produto fosse eticamente defensável, respeitando o Código da ACM, princípios de engenharia e a integridade profissional.

Referências:

- Moor, J. H. (1985).
What is computer ethics? *Metaphilosophy*, 16(4), 266–275.
- Huff, C. (1990).
Unintentional power in computer systems. *Computer Ethics and Professional Responsibility* (Livro 1, Cap. 4, p. 49).
- Bynum, T. W. (2001).
Computer and information ethics. *The Stanford Encyclopedia of Philosophy*.
- Gotterbarn, D. (1991).
Informatics and professional responsibility. *Computing and Philosophy Conference*.
- Association for Computing Machinery. (2018).
ACM Code of Ethics and Professional Conduct <https://www.acm.org/code-of-ethics>
- Ética para Engenheiros. (Material da UC, p. 47; p. 57–60).