

Desenvolvimento e Aplicação de uma Metodologia para Estimativa e Interpretação da Proficiência dos Alunos da Estácio utilizando TRI e Inteligência Artificial

Letícia Correa¹

Corporativo YDUQS
Rio de Janeiro, RJ, Brasil
leticia.correa@yduqs.com.br

Luís Claudio Saldanha²

Estácio
Rio de Janeiro, RJ, Brasil
luis.dallier@ensineme.com.br

Mariana Soares³

Psicologia³
Campo Grande, MS, Brasil
mariana.soares@ufms.br

Tufi Soares⁴

Estatística, UFJF⁴
Juiz de Fora, MG, Brasil
tufi.soares@ufjf.br

Resumo

O ensino a distância (EAD) exige adaptações específicas nos métodos de avaliação para garantir sua eficácia. Este estudo apresenta uma metodologia de cálculo da proficiência dos alunos da Estácio, integrando a Teoria de Resposta ao Item (TRI) de três parâmetros e inteligência artificial (IA). Foram analisadas 9.905.244 respostas referentes a 1.553 disciplinas em 2024, utilizando R e Python para estimação dos parâmetros psicométricos dos itens. Escalas de proficiência padronizadas de 0 a 100 foram construídas, classificando os alunos em cinco faixas de desempenho. A metodologia foi aplicada a cerca de 1 milhão de alunos-disciplina, com resultados que demonstraram boa distribuição nas faixas de proficiência e melhor desempenho dos alunos no EAD. Provas equilibradas automatizadas aumentaram a validade das avaliações e reduziram custos. Futuramente, a plataforma será expandida para oferecer recomendações personalizadas de estudo com base na análise psicométrica individual, potencializando o desenvolvimento das habilidades dos estudantes.

Palavras-chave: TRI, avaliação, proficiência, graduação.

Os autores agradecem a ajuda do prof. Sergio Silva e Tiago Figueredo pelas reflexões pedagógicas.

1 Introdução

O ensino a distância (EAD) tem se consolidado como uma modalidade fundamental na educação contemporânea, especialmente após a pandemia, quando a necessidade de alternativas ao ensino presencial se intensificou. Entre suas principais vantagens destacam-se o menor custo e a flexibilidade de horários, características que o tornam atrativo para uma ampla gama de estudantes¹. No entanto, a simples transposição do modelo pedagógico presencial para o ambiente digital não garante a eficácia do ensino a distância. Para que o EAD seja verdadeiramente eficaz, são necessárias adaptações metodológicas e a reavaliação de práticas fundamentais, como os processos de avaliação educacional.

A adaptação do modelo avaliativo é um dos pilares essenciais nesse contexto. No ensino presencial, o contato direto entre professor e aluno permite acompanhar de forma contínua o progresso da turma, possibilitando ajustes pontuais nas avaliações conforme o desempenho coletivo. No EAD, essa dinâmica torna-se mais complexa, uma vez que a ausência da interação presencial dificulta a percepção direta do docente sobre o desenvolvimento dos alunos. Diante desse desafio, surge a seguinte questão: como adaptar o sistema de avaliação para oferecer uma análise precisa e individualizada da aprendizagem no ensino a distância?

Este projeto foi concebido a partir dessa necessidade de aprimorar as práticas avaliativas, oferecendo uma abordagem mais rigorosa, adaptativa e personalizada. A solução proposta integra a Teoria de Resposta ao Item (TRI) de três parâmetros^{2,3} — que estima a dificuldade (b), a discriminação (a) e o acerto casual (c) dos itens — com modelos de inteligência artificial (IA), permitindo uma construção mais precisa e balanceada das avaliações. Além disso, a aplicação da TRI possibilita o cálculo da proficiência individual dos alunos, fornecendo uma estimativa padronizada e comparável de suas habilidades em cada disciplina. Essa mensuração da proficiência transcende as limitações das notas tradicionais, oferecendo um indicador mais fiel do conhecimento efetivamente adquirido.

Assim, a utilização conjunta da TRI e da IA não apenas aprimora a seleção e calibração dos itens avaliativos, mas também permite monitorar de forma contínua o desenvolvimento dos alunos no EAD, com a geração de dados que subsidiam intervenções pedagógicas mais eficazes e a personalização dos percursos de aprendizagem.

2 Metodologia

Para a realização deste estudo, foi utilizada a base de dados da Estácio, composta por avaliações supervisionadas (AV e AVS) que cobrem integralmente o conteúdo de diversas disciplinas de graduação. A coleta de dados ocorreu no ano de 2024, resultando em um total de 17M de respostas, distribuídas entre 3.594 disciplinas. Cada avaliação é composta por provas de múltipla escolha, com 10 questões, cada uma apresentando cinco alternativas. As questões utilizadas foram extraídas do banco de questões da Estácio, desenvolvido por especialistas em cada área, com base em matrizes de conteúdo validadas pelo Ministério da Educação (MEC).

A análise dos dados foi conduzida utilizando as linguagens de programação R e Python. Para a estimação dos parâmetros psicométricos dos itens e cálculo da proficiência do aluno, aplicou-se a Teoria de Resposta ao Item (TRI), utilizando a biblioteca Mirt⁴. Essa biblioteca é especializada na estimação de parâmetros de itens dicotômicos, seguindo o modelo de três parâmetros, que considera a dificuldade, a discriminação e o acerto casual de cada item.

Inicialmente, os parâmetros dos itens foram estimados utilizando o modelo logístico de três parâmetros. Os parâmetros psicométricos de todos os itens foram calculados para cada uma das amostras disponíveis, adotando como critérios mínimos: 20 respostas por item, 200 respostas por disciplina e mais de 5 itens por disciplina. Em seguida, os itens foram classificados segundo sua dificuldade (β) e qualidade (discriminação, α). Itens com dificuldade (β) inferior a -0,7 foram classificados como fáceis, enquanto aqueles com β superior a 0,6 foram classificados como difíceis. Quanto à discriminação (α), os itens foram classificados como: qualidade muito baixa (<0,5), baixa (0,5–1,2), média (1,2–2,0) e alta (>2,0).

Após a calibração dos itens, procedeu-se à construção da escala de proficiência para cada disciplina. A escala foi definida individualmente para cada disciplina, baseada na população fixa de alunos matriculados em 2024, considerando apenas disciplinas com mais de 200 alunos. As escalas foram padronizadas em um intervalo de 0 a 100 e divididas em cinco faixas de proficiência: iniciante, elementar, básico, elevado e avançado. Esta divisão permitiu uma melhor diferenciação dos níveis de aprendizagem dos alunos. Para a fixação da escala, foram utilizados apenas itens calibrados que atendessem aos critérios de quantidade de respostas e qualidade psicométrica, assegurando sua estabilidade para aplicações futuras.

O cálculo da proficiência dos alunos foi realizado com base no desempenho nas provas compostas por itens calibrados. Onde se exigiu o mínimo de oito questões calibradas respondidas. A proficiência do aluno foi estimada a partir do padrão de acertos e erros nas questões respondidas, considerando as características psicométricas dos itens, conforme o modelo logístico de três parâmetros. A pontuação final representa a habilidade mais provável do aluno na disciplina avaliada.

Além disso, uma API integrou o sistema de cálculo da TRI ao sistema de provas. No sistema SIRIUS, onde todos os itens são armazenados e as provas são administradas, foi implementada uma inteligência artificial (IA) desenvolvida em Node.js para automatizar a montagem das provas. A IA organiza as questões em ordem crescente de dificuldade, partindo das mais fáceis até as mais difíceis. O setor de ensino determina previamente a quantidade de questões por nível de dificuldade, sendo obrigatória a inclusão de duas questões não calibradas para renovação do banco. Essas avaliações são denominadas provas equilibradas.

Após a realização da prova pelo aluno, a proficiência é automaticamente calculada e armazenada no sistema, alimentando dashboards acadêmicos e subsidiando intervenções pedagógicas e ações de melhoria no processo de ensino-aprendizagem. A proficiência foi dividida de 0-100+ e dividida em 5 faixas, sendo: iniciante, elementar, básico, elevado e avançado. Cada disciplina possui uma faixa de pontos de acordo com sua escala.

3 Resultados

O processo de análise permitiu a obtenção de medidas de proficiência para uma ampla base de alunos e disciplinas da Estácio. Ao longo do ano de 2024, foram estimadas proficiências para aproximadamente 1 milhão de alunos-disciplina e 265 mil alunos únicos que realizaram provas digitais calibradas. Abrangendo um total de 970 disciplinas em diversos cursos. A Figura 1 apresenta a distribuição geral das proficiências dos alunos ao longo das faixas previamente definidas:

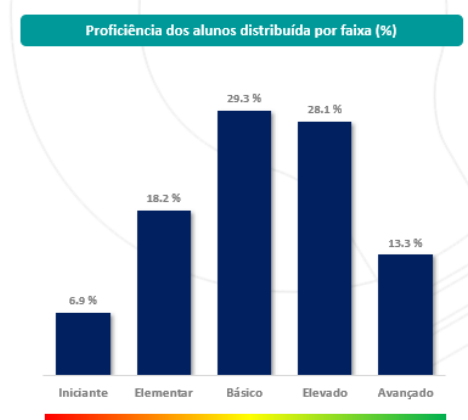


Figura 1: Distribuição total das proficiências dos alunos dentro das faixas de proficiência previamente definidas.

Adicionalmente, foi realizada uma análise comparativa entre a nota bruta (pontuação) e a proficiência estimada, com o objetivo de evidenciar diferenças em função da modalidade de ensino (Figura 2) e das áreas de conhecimento (Figura 3).

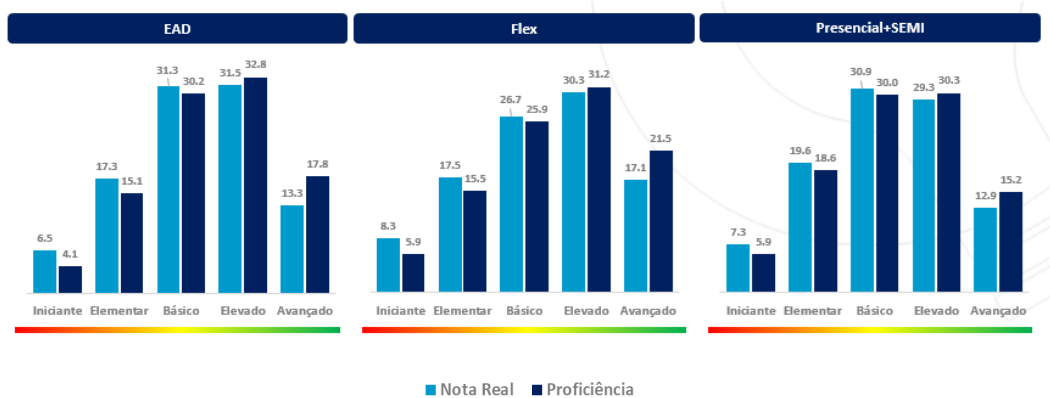


Figura 2: Comparação entre a classificação dos alunos por faixa de nota bruta e faixa de proficiência no período de 2024.3, segmentada por modalidade de ensino (presencial e EAD).

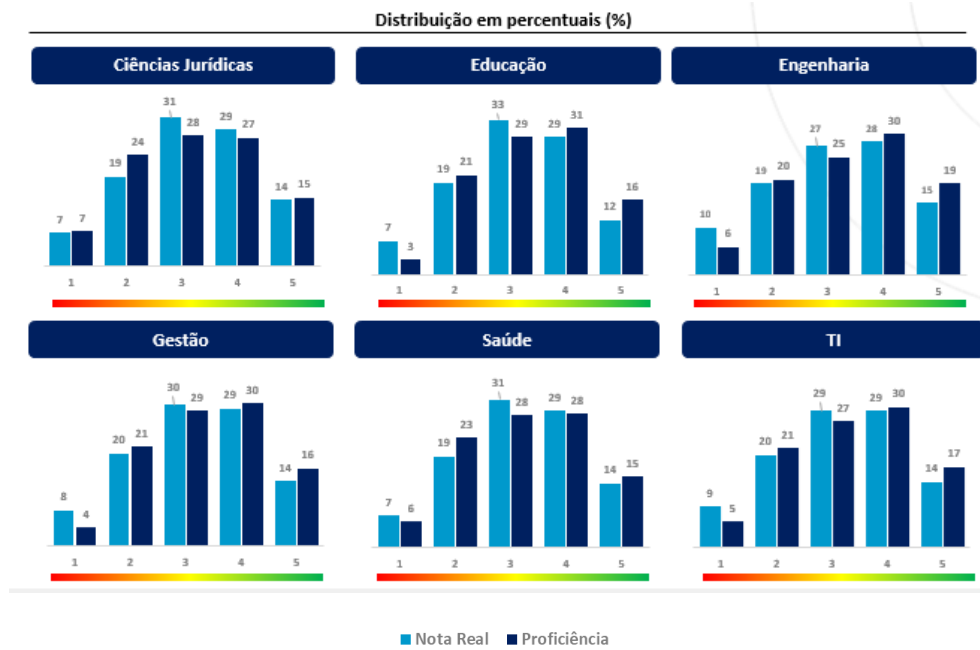


Figura 3: : Comparação entre a classificação dos alunos por faixa de nota bruta e faixa de proficiência no período de 2024.3, segmentada por área de conhecimento.

3.1 Discussão

A distribuição dos alunos nas faixas de proficiência revelou uma dispersão estatisticamente equilibrada, com leve assimetria à direita, o que indica um desempenho geral positivo. Ao estratificar por modalidade de ensino, observa-se que os estudantes da modalidade a distância (EAD) apresentaram, em média, níveis mais elevados de proficiência. No ensino presencial, cerca de 25% dos alunos concentram-se nas faixas iniciais, sugerindo maiores desafios de aprendizagem nesse grupo.

A análise por área de conhecimento revelou padrões distintos conforme a dificuldade das provas. Considera-se que as avaliações foram mais fáceis quando a proficiência média dos alunos superou suas respectivas notas — como observado nos cursos de Ciências Jurídicas — o que sugere que os estudantes apresentaram maior habilidade latente do que a pontuação obtida indica. Em contrapartida, em áreas como Engenharia e Tecnologia da Informação, onde as notas médias superaram os níveis de proficiência estimados, infere-se maior dificuldade nas avaliações, possivelmente exigindo estratégias de resolução mais automatizadas ou menos aderentes à habilidade medida. Esses resultados destacam a capacidade da Teoria de Resposta ao Item em oferecer diagnósticos mais sensíveis e acurados sobre o desempenho estudantil.

Esses achados reforçam a sensibilidade da Teoria de Resposta ao Item em captar com maior precisão as competências dos estudantes, contribuindo para diagnósticos mais robustos e para a formulação de práticas pedagógicas mais eficazes.

4 Conclusões e Considerações Finais

Os resultados deste estudo ressaltam o impacto positivo da integração entre a Teoria de Resposta ao Item (TRI) e a inteligência artificial (IA) no processo de avaliação educacional da Estácio. A implementação da plataforma inteligente de avaliação, associada ao cálculo da proficiência dos alunos, contribuiu para uma melhoria significativa na precisão, validade e qualidade das avaliações. A utilização da TRI permitiu mensurar com maior fidelidade a habilidade real dos alunos, independentemente da dificuldade das provas, promovendo maior equidade entre os avaliados.

Adicionalmente, a automação dos processos de montagem de provas, baseada na classificação psicométrica das questões, gerou ganhos relevantes em termos de eficiência operacional e redução de custos. A introdução de provas equilibradas — organizadas conforme o nível de dificuldade das questões — garantiu a padronização da complexidade das avaliações, assegurando a comparabilidade dos resultados entre diferentes turmas e períodos.

Em relação às perspectivas futuras, o projeto visa expandir as funcionalidades da plataforma com foco ainda maior no desenvolvimento da proficiência dos alunos. Uma das principais evoluções planejadas é a implementação de mecanismos de recomendação personalizada de estudos, baseados na análise individualizada do desempenho de cada aluno. Utilizando a robustez dos dados psicométricos da TRI e a capacidade preditiva da inteligência artificial, será possível identificar com maior precisão as áreas de dificuldade específicas de cada estudante. A partir disso, a plataforma poderá oferecer sugestões direcionadas de conteúdos, exercícios e atividades complementares, promovendo um ciclo contínuo de aprendizagem e fortalecimento de habilidades.

5 Referências

1. Kalaki, C. J. et al. O avanço do ensino remoto, o crescimento do ensino a distância e a importância do profissional de tutoria. *Research, Society and Development* 13, e6713144779 (2024).
2. Hambleton, R. K. ., Swaminathan, Hariharan. & Rogers, H. Jane. *Fundamentals of Item Response Theory*. (SAGE Publications, Inc, 2012).
3. Francisco De Andrade, D., Tavares, H. R. & Da Cunha Valle, R. *Teoria Da Resposta Ao Item: Conceitos e Aplicações*.
4. Chalmers, R. P. Mirt: A multidimensional item response theory package for the R environment. *J Stat Softw* 48, (2012).