

Automação na correção de provas com Python e hospedagem web: Ciência de dados aplicada à educação

Automation in exam grading with Python and web hosting: data science applied to education

Paulo Henrique Cerutti *(ORCID 0000-0001-6664-8449), **Luan Tiago dos Santos Carbonari** (ORCID 0000-0002-7797-5194), **Carlos Zacarias Joaquim Junior** (ORCID 0000-0001-8681-9946), **Mauro Bitencourt de Souza** (0009-0003-8723-6285), **Henrique de Sá Albino** (0009-0002-9654-7539), **Jefferson Luís Meirelles Coimbra** (ORCID 0000-0001-9492-6055)

Universidade do Estado de Santa Catarina, Lages, SC, Brasil. *Autor para correspondência:
paulohcerutti@gmail.com

Submissão: 01 de Novembro, 2024 | Aceite: 15 de Março, 2025

RESUMO

A ciência de dados utiliza métodos estatísticos e ferramentas computacionais para extrair conhecimento a partir de dados ou informações sobre um determinado assunto de forma automatizada. No contexto educacional a automação da correção de provas proporciona maior eficiência e precisão, sendo especialmente relevante nas ciências sociais e agrárias, onde o volume de alunos, de questões e a necessidade de correção rápida são eventos frequentes. Assim, o objetivo deste trabalho foi desenvolver um sistema de automação chamado *AutoCorrect*, utilizando *Python* e hospedagem Web para a correção de provas acadêmicas. O mecanismo permite a geração de provas personalizadas, com gabaritos aleatórios, e realiza a correção automática das respostas submetidas, otimizando significativamente o processo de avaliação. A metodologia incluiu o uso de bibliotecas específicas do *Python*, como *Pandas*, *Numpy* e *Matplotlib*, para o processamento de dados, análise estatística e visualização gráfica. Para tanto, foi realizada uma aplicação prática em turmas dos cursos de Agronomia e Medicina veterinária, nas quais os alunos responderam às provas geradas. As respostas dos alunos foram comparadas diretamente com o gabarito fornecido pelos professores, o que resultou na geração automática das pontuações de cada estudante, baseadas em seus acertos e erros. Além disso, o sistema permitiu a análise detalhada das notas por meio de medidas descritivas, como média, desvio padrão da média, variância e assimetria, oferecendo uma visão clara do desempenho da turma. Gráficos complementares foram gerados para facilitar a visualização da distribuição das notas e do desempenho individual em relação à média geral. A integração da ciência de dados com ferramentas computacionais como *Python* e hospedagem web, não apenas otimizam o processo de correção de provas, mas também fornecem maior clareza e agilidade à análise do desempenho dos alunos. O sistema *AutoCorrect* apresenta-se como uma ferramenta eficiente e inovadora para várias áreas da educação, especialmente nas ciências agrárias e sociais.

PALAVRAS-CHAVE: *AutoCorrect*. Análise de dados. Aprendizagem de Máquinas. Ferramentas Computacionais. Relatórios Educacionais.

ABSTRACT

Data science employs statistical methods and computational tools to automatically extract knowledge from data or information on a specific subject. In the educational context, automating exam grading enhances efficiency and accuracy, which is especially relevant in social and agrarian sciences, where a high volume of students, questions, and the need for quick grading are common. Thus, the objective of this study was to develop an automation system called *AutoCorrect*, utilizing *Python* and web hosting for the grading of academic exams. The mechanism allows for the generation of customized exams with randomized answer keys and performs automatic grading of submitted responses, significantly optimizing the evaluation process. The methodology included the use of specific *Python* libraries, such as *Pandas*, *Numpy*, and *Matplotlib*, for data processing, statistical analysis, and graphical visualization. A practical application was conducted in Agronomy and Veterinary Medicine classes, in which students took the generated exams.

Publisher's Note: UDESC stays neutral concerning jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Student responses were directly compared to the answer key provided by instructors, resulting in the automatic generation of scores for each student based on correct and incorrect answers. Additionally, the system enabled a detailed analysis of grades through descriptive measures, such as mean, mean standard deviation, variance, and skewness, providing a clear view of class performance. Supplementary charts were generated to facilitate visualization of the grade distribution and individual performance relative to the overall average. The integration of data science with computational tools like *Python* and web hosting not only optimizes the grading process but also provides greater clarity and agility in analyzing student performance. The *AutoCorrect* system proves to be an efficient and innovative tool for various educational fields, especially in agrarian and social sciences.

KEYWORDS: *AutoCorrect*. Data Analysis. Machine Learning. Computational Tools. Educational Reports.

INTRODUÇÃO

A ciência de dados tem se consolidado como uma ferramenta essencial para a educação, proporcionando soluções que otimizam processos de ensino e avaliação. Em áreas como Agronomia e Medicina Veterinária, a aplicação de técnicas de ciência de dados permite automatizar tarefas repetitivas, como a correção de provas, tornando a avaliação mais rápida, precisa e eficiente.

Esse processo vai além da simples análise de grandes volumes de dados ("Big Data"), pois envolve principalmente a criação de soluções práticas para desafios reais, tanto dos educadores quanto dos alunos. Ao utilizar algoritmos de aprendizado de máquina e técnicas estatísticas, a ciência de dados transforma informações em ações concretas, facilitando o trabalho docente e aprimorando a experiência educacional (VAN DER AALST 2016).

A pandemia de COVID-19 destacou ainda mais a necessidade de soluções automatizadas no ambiente educacional. Os educadores nas áreas de ciências agrárias em particular, se depararam com a importância de utilizar tecnologias que facilitem a correção de avaliações e a análise de desempenho dos estudantes. Na área da medicina veterinária, autores destacaram a importância de se desenvolver plataformas digitais para disseminar a didática e aprendizado durante a pandemia, observando que este modelo é auxiliar e nunca substitutivo de formas clássicas de ensino (MASSARI et al. 2022).

Com a automatização de correções, a ciência de dados exemplifica a sua capacidade de transformar dados brutos em soluções eficazes, proporcionando uma avaliação mais ágil e precisa (ALDRIYE et al. 2019, KASINATHAN et al. 2022, LIGUORI & WINKLER 2020). Na automação da correção de provas, por exemplo, a coleta de dados (respostas dos alunos para testes aplicados), é seguida por um processo de preparação, que inclui a organização e o pré-processamento dos dados. Algoritmos são aplicados para identificar padrões nas respostas dos acadêmicos e automatizar a correção, fornecendo feedbacks consistentes e precisos. Essa abordagem não só economiza tempo, mas também garante uma avaliação mais justa e padronizada para todos os alunos (JAMES et al. 2023).

Neste sentido, vários sites e aplicativos já são utilizados para a correção automatizada de provas, como Kahoot, Google Forms, ExamSoft, ZipGrade e Moodle, dentre outros. Por exemplo, o ExamSoft é voltado para avaliações seguras, bloqueando o acesso à internet durante os exames, enquanto o ZipGrade transforma

dispositivos móveis em scanners para a correção rápida de avaliações de múltipla escolha. O Moodle oferece um ambiente virtual de aprendizagem de código aberto, usado em algumas instituições, porém a curva de aprendizagem para novos usuários sem o conhecimento em *Learning Management System* (sistema de gerenciamento de aprendizagem) pode ser difícil; a manutenção do Moodle pode exigir conhecimento técnico para trabalhar com servidores, atualizações e segurança; dependendo do número de usuários e das funcionalidades ativadas, o Moodle pode exigir bastante poder de processamento e largura de banda do servidor utilizado (MUSTAFA & ALI 2023, RODILLAS et al. 2023).

No entanto, muitas dessas plataformas direcionam o seu foco em um escopo mais amplo de atividades educacionais e não foram essencialmente desenvolvidas para a correção de provas, o que limita sua eficiência e eficácia. Com base nisso, o sistema proposto neste trabalho diferencia-se ao enfatizar diretamente a correção automatizada de provas, sendo testado nos cursos de Agronomia e Medicina Veterinária, com flexibilização de expansão para todas as áreas da ciência. O mesmo utiliza um código em linguagem *Python* com funções específicas para organizar e corrigir as respostas dos alunos de forma eficaz.

Além disso, a criação de uma plataforma web com a funcionalidade “Multclick” oferece uma interface simples e intuitiva, facilitando o acesso dos docentes à correção automatizada, fornecendo resultados rápidos e precisos. Por exemplo, considera-se uma sala de aula com 40 alunos matriculados em determinada disciplina e aptos a realizar uma prova. Ao aplicar uma prova com 10 questões, sendo que cada questão tem 10 itens, tem-se $40 * 10 * 10 = 4000$ itens a serem corrigidos.

Supondo que o professor aplique quatro provas no semestre, tem-se 16000 itens para correção. Se o mesmo demorar 30 segundos para corrigir cada item, totaliza cerca de 480000 segundos destinados a esta tarefa. Isso gera aproximadamente 133 horas de trabalho, ou seis dias.

Com o *AutoCorrect*, após a confecção dos arquivos necessários, aplicação da prova e coleta das respostas dos alunos, a correção é realizada em poucos minutos, com a opção do aluno mesmo preencher as suas respostas e enviar ao sistema. Além disso, o professor pode dispor de um banco de questões de sua disciplina, mediante o número de questões desejado, o sistema irá sortear a quantidade de perguntas em cada avaliação. Este fato amplia as possibilidades de realização de provas e diminui drasticamente a conduta de plágio entre os discentes.

A solução apresentada demonstra a importância de integrar ciência de dados à educação, oferecendo uma proposta concreta às demandas específicas do processo avaliativo no ensino superior (ZHANG et al. 2022).

A partir de então, os objetivos deste trabalho foram: desenvolver um código em linguagem de programação *Python* e a criação de um ambiente de internet na forma de Hospedagem Web, com a função Multclick (clique e obtenha o resultado), destinados a correção automatizada das provas, de modo a facilitar o trabalho dos docentes e garantir uma avaliação justa, eficiente e precisa aos alunos.

MATERIAL E MÉTODOS

A correção automatizada de provas pode ser executada em dois ambientes:

Python e hospedagem Web.

Bibliotecas Utilizadas

Diversas bibliotecas *Python* foram empregadas no código para realizar as análises e processamento de dados. As bibliotecas mais relevantes foram destacadas na Tabela 1.

Tabela 1. Descrição das principais bibliotecas utilizadas na confecção do sistema de correção de provas.

Table 1. Description of main libraries used creation of the test correction system.

Biblioteca	Função
Pandas	Manipulação e análise de dados, incluindo leitura e escrita de arquivos em formatos diversos (csv; xlsx, dentre outros).
NumPy	Operações matemáticas e manipulação de arrays (estruturas de dados que contêm elementos de um mesmo tipo, organizados de forma contígua na memória).
Matplotlib e Seaborn	Visualização de dados e criação de gráficos.
Scikit-learn	Aplicação de algoritmos de aprendizado de máquina e validação de modelos.
<u>OpenPyXL</u> ou <u>xlrd</u>	Leitura e escrita de arquivos Excel.

Fonte: HELMUS & COLLINS 2016.

Estas bibliotecas foram instaladas no ambiente virtual do Anaconda com o comando *conda install nome_da_biblioteca* ou, em alguns casos, usando o comando *pip install nome_da_biblioteca*, para instalar pacotes que não estão disponíveis diretamente através do Anaconda (ROLON-MÉRETTE D et al. 2016).

Utilização de Arquivos Locais

Ao considerar tanto o código desenvolvido em *Python*, quanto a hospedagem Web, ambos fazem uso de arquivos armazenados no computador particular. Estes arquivos podem incluir conjuntos de dados em formatos como: .txt; .csv; .tsv; .xls; .xlsx; .json; .jpg; .png; .zp, dentre outros. Os arquivos utilizados foram acessados e manipulados diretamente a partir do diretório de trabalho do *Jupyter Notebook*. O caminho para esses arquivos foi especificado no código para garantir que os dados fossem lidos e processados corretamente (a partir de pastas no computador pessoal), de acordo com as disciplinas de melhoramento genético vegetal e estatística básica, dos cursos de Agronomia e Medicina Veterinária, respectivamente.

Para a execução deste código destinado a correção de provas e para o uso da hospedagem Web, foi necessária a confecção de alguns arquivos. Estes são diretamente relacionados com os dados de cada professor, e devem ser salvos em uma pasta diretório do computador particular do usuário. Os arquivos são: *i)* Nome dos alunos: documento com a lista dos alunos matriculados na disciplina, aptos a realizar a prova. Este documento pode ser em formato .xlsx; *ii)* Documento com as questões da prova: em formato .docx; *iii)* Gabarito do professor: documento com as repostas das questões escolhidas no arquivo de provas, em formato .xlsx; *iv)* Peso de cada questão: arquivo que indica o peso atribuído a cada questão, calculado inclusive

pela quantidade de itens por questão, no formato .xlsx. Nas figuras 1 a 4, são demonstradas as formas de confecção dos arquivos, utilizando nomes e e-mails fictícios, de modo a preservar a identificação dos alunos. Estes arquivos são essenciais para o início do processo.

A	B	C
Index	Nome	Email
1	Ana Clara	ana.clara@example.com
2	Pedro Henrique	pedro.henrique@example.com
3	Mariana Souza	mariana.souza@example.com
4	João Vitor	joao.vitor@example.com
5	Laura Mendes	laura.mendes@example.com
6	Gabriel Oliveira	gabriel.oliveira@example.com
7	Beatriz Santos	beatriz.santos@example.com
8	Rafael Lima	rafael.lima@example.com
9	Júlia Ferreira	julia.ferreira@example.com
10	Lucas Pereira	luca.pereira@example.com

Q1. Os dados a seguir representam o rendimento de grãos de parcelas de soja, em gramas. 120; 100; 150; 250; 250; 155; 175; 185; 190; 200; 210; 215; 220. Pergunta-se: qual a escala de medida adotada? Qual o rendimento médio de grãos? A variação em g do conjunto? A conversão desta variação em percentual? Uma estimativa de erro na amostragem? Com base nisso, responda assinalando no gabarito apenas a(as) alternativa(s) correta(s): (1) Ordinal; 187; 45,74; 25% e 13. (2) Racional; 187; 45,74; 25% e 13; (3) Intervalar; 186,15; 45,74; 24,57% e 12,68; (4) Racional e Ordinal; 187; 45,90; 23,9% e 15; (5) Racional; 186,15; 45,74; 24,57% e 12,68; (6) Nenhuma das anteriores.

Q2. Sobre medidas para dados quantitativos, escalas de medida e tabela de frequência, assinale a/as alternativas corretas: (1) Mesmo um especialista em uma área de pesquisa pode não ser capaz de obter uma amostra que seja verdadeiramente representativa, então é melhor escolher uma amostra aleatória; (2) o objetivo da estatística é fazer inferências sobre uma população baseadas na informação de amostras; (3) Distribuições de frequência são sempre simétricas; (4) A escala ordinal é algumas vezes usada, mesmo que informação numérica mais precisa seja disponível; (5) Pressão barométrica é usualmente registrada em escala ordinal; (6) Operações aritméticas podem ser efetuadas sobre escalas nominais e de intervalo, mas não podem ser aplicadas a números de escalas de razão e ordinal; (7) A estatística provê estimativas de parâmetros.

2

A	B	C	D	E	F	G	H	I	J	K
QUESTAO	1	2	3	4	5	6	7	8	9	
Q1					5	0.2525				
Q2	X	X		X			X			
Q3	1	2	3		5		7			

3

A	B	C	D	E
Questão	peso	Subitens		
Q1	3	2		
Q2	3	4		
Q3	4	5		

4

Figuras 1 a 4. Formação dos arquivos para a elaboração completa do processo de correção de provas, sendo: 1- arquivo com o nome e e-mail dos alunos; 2- arquivo com as questões da prova; 3- arquivo com a/as respostas corretas por questão e 4- peso atribuído a cada questão.

Figures 1 to 4. Creation of files for the complete preparation test correction process, being: 1- file with the name and email of the students; 2- file with the test questions; 3- file with the correct answers per question and 4- weight assigned to each question.

Partes do código

O código em *Python* para a automação na correção de provas foi formulado em partes: sendo elas: *i*) Aleatorização de gabaritos conforme os nomes dos alunos listados; *ii*) Geração de questões aleatorizadas, de acordo com os gabaritos aleatorizados; *iii*) Compilação de arquivos de provas aleatorizadas para impressão, conferência e entrega aos alunos no dia da prova; *iv*) Correção das provas, de acordo com as respostas assinaladas nos gabaritos aleatórios. Esta correção é realizada tanto pela marcação de “x” no gabarito, bem como a escolha de números correspondentes aos itens corretos (Ex: questão 3: itens corretos 1, 2, 3, 5 e 7) e também pela marcação de valores resultantes de cálculos no gabarito (Ex: questão 1 das figuras representadas anteriormente, onde a resposta correta para o item 6 dessa questão é 0,2525; *v*) correção com upload do arquivo peso das questões. Neste item, se informa o peso de cada questão em um arquivo para a correção; *vi*) Geração de um mapa de calor. Comando usado para confeccionar um gráfico de calor, com base nas pontuações dos alunos e *vii*) Confecção de um gráfico com desvio padrão em relação à média geral.

Uma vez configurado o ambiente e instaladas as bibliotecas necessárias, o código foi executado no *Jupyter Notebook*. Todas as partes do código se encontram em um endereço do google drive, na pasta denominada “Correção Provas *Python*”, disponível para download por qualquer usuário:

https://drive.google.com/drive/folders/1xUYkB7EsXo-F9CXleLhZ0wrgK-SN3wtH?usp=drive_link.

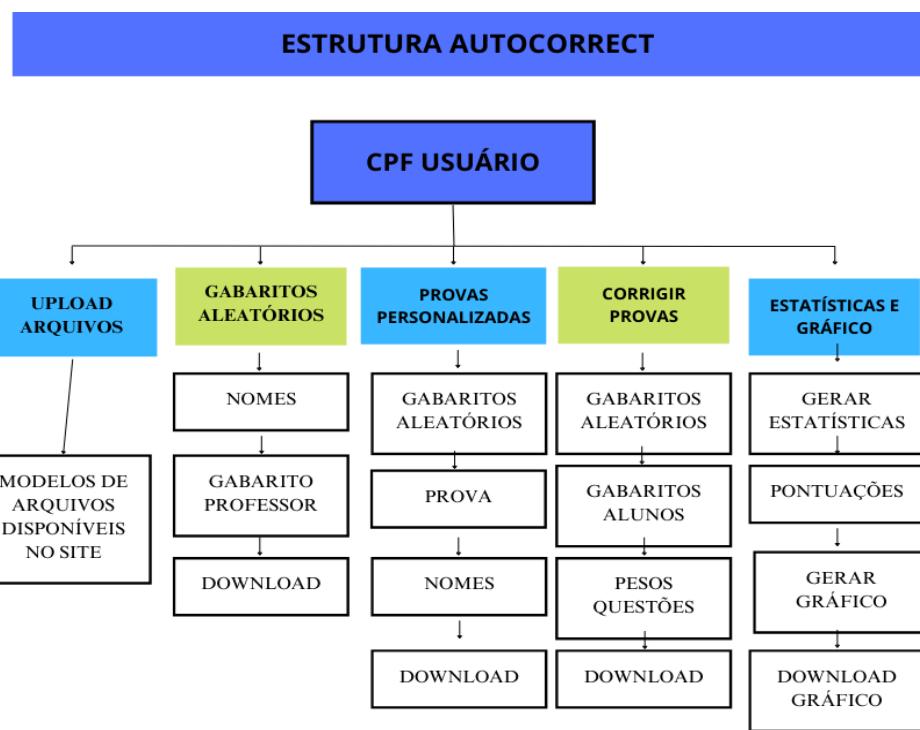
Hospedagem Web

O site denominado *AutoCorrect* foi projetado basicamente para garantir uma opção Multiclick aos usuários deste sistema de correção de provas automatizadas. O mesmo pode ser acessado pelo endereço: <http://coimbrajefferson.pythonanywhere.com>.

O *AutoCorrect* apresenta acesso aberto, contudo os autores detêm a patente e o registro da marca de programas para computador. Para acessar, o CPF deve ser escrito sem pontos, apenas com números, exemplo:11111111111. A escolha do CPF como forma de entrada se deve pelo fato de cada pessoa conter uma identificação específica.

Porém, salientamos que se os usuários definirem a utilização de CPFs fictícios e por acaso utilizarem o mesmo número (11111111111), o sistema não funcionará corretamente, em virtude de que cada professor possui seus próprios arquivos.

O funcionamento do site *AutoCorrect*, segue o seguinte fluxograma:



Cada uma das partes tem a seguinte visão geral:



Figura 5. Visão geral do acesso ao site para a correção automatizada das provas e interface inicial do site para Upload dos arquivos necessários na realização do processo.

Figure 5. Overview of access to the website for automated correction and the website's initial interface for uploading the files required to complete the process.

Após o acesso, o usuário se depara com a interface do site (Figuras 5 a 9), sendo possível realizar o Upload dos arquivos para o processo automatizado de correção de provas, de acordo com a demanda de cada professor.

Gerar Gabaritos Aleatorizados

CPF:

Caminho do arquivo com os nomes dos alunos (XLSX):
Gerado automaticamente

Caminho do arquivo Excel com o gabarito do professor:
Gerado automaticamente

Número de questões:

Seed fixa (opcional):

Usar seed fixa?

Gerar Gabaritos

Figura 6. Fração do site diretamente responsável pela formação dos gabaritos aleatorizados para cada aluno.

Figure 6. Fraction of the site directly responsible for creating randomized templates for each student.

Gerar Provas Personalizadas

CPF:

Caminho do arquivo Excel com os gabaritos aleatorizados:
Gerado automaticamente

Caminho do arquivo DOCX com as questões:
Gerado automaticamente

Caminho do arquivo Xlsx com os nomes dos alunos:
Gerado automaticamente

Caminho para salvar as provas personalizadas:
Gerado automaticamente

Gerar Provas

Unir Provas

Caminho dos arquivos que serão unidos:
/home/coimbrajefferson/CorrecPro/uploads/1111111111/provas_geradas/

Unir Provas

Figura 7. Frações responsáveis por gerar as provas para cada pessoa e união destas provas em único documento Microsoft Word para impressão.

Figure 7. Fractions responsible for generating the tests for each person and combining these into a single Microsoft Word document for printing.

As provas para cada estudante são confeccionadas de acordo com o gabarito aleatorizado gerado para cada estudante.

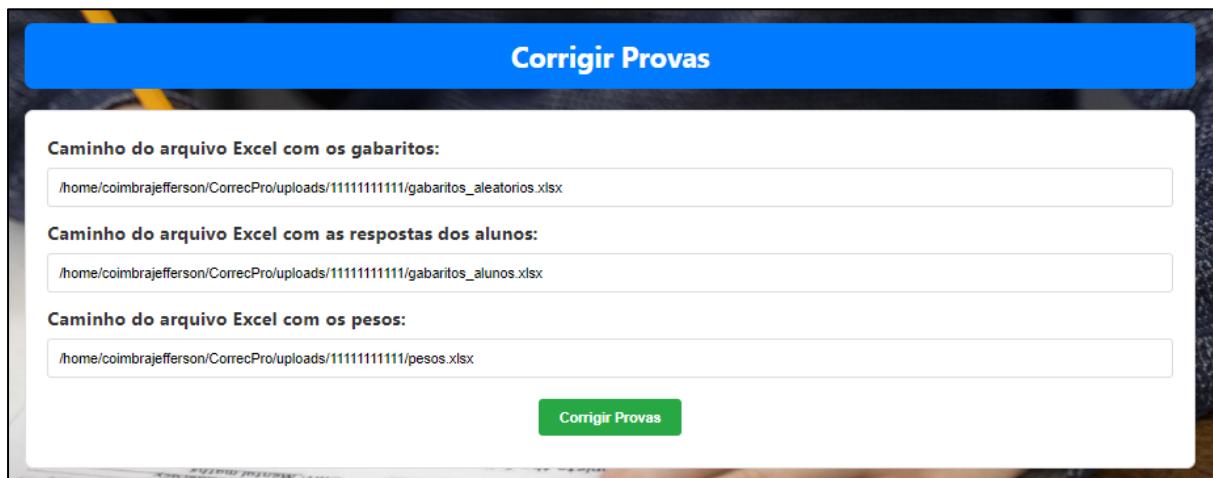


Figura 8. Parte do site destinada a correção de provas e possibilidade de download dos resultados.
Figure 8. Part of the website intended for correcting tests and the possibility of downloading results.

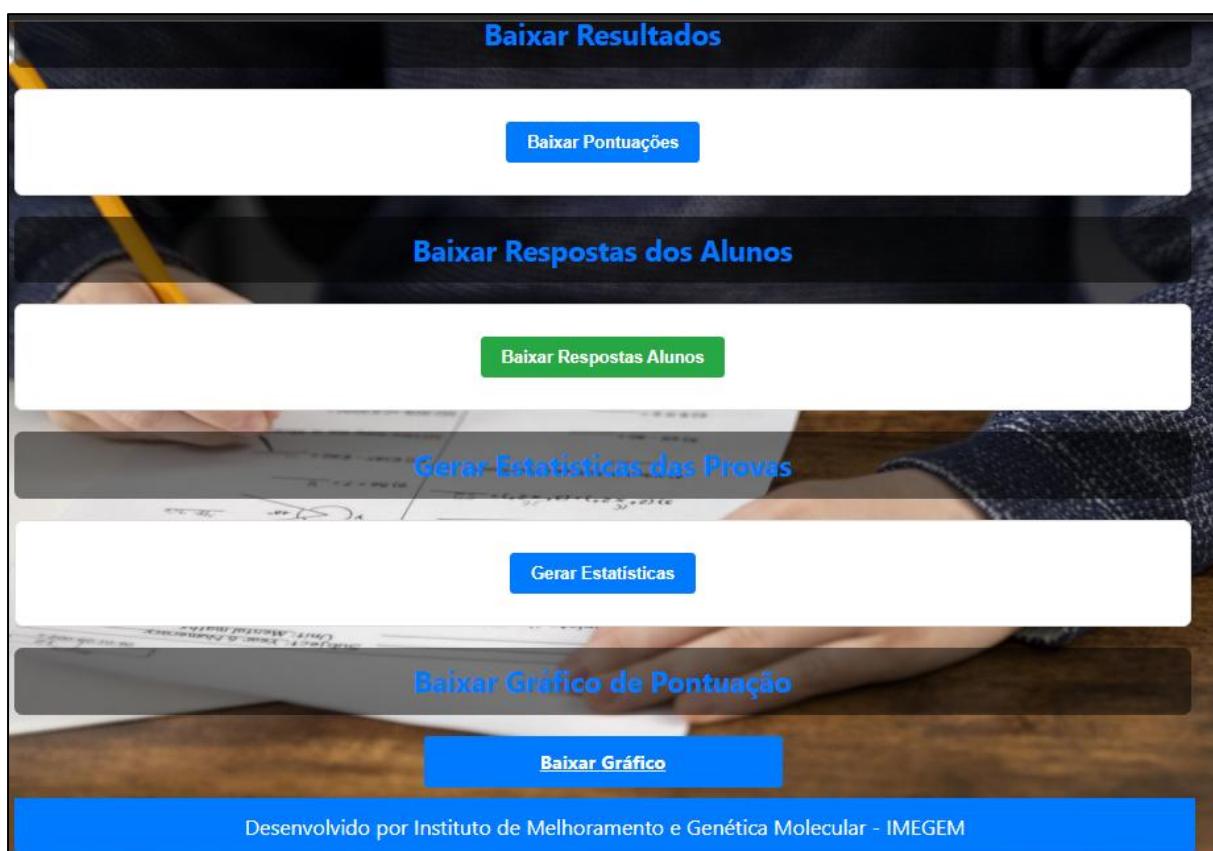


Figura 9. Parte do site destinada ao download das respostas dos alunos, da geração de estatística descritiva e gráfico com as pontuações.
Figure 9. Part of the website intended for downloading students' responses, generating descriptive statistics and a graph with scores.

Considerações sobre o funcionamento do *AutoCorrect*: Na mesma pasta do google drive, foi disponibilizado um arquivo pdf com as instruções para uso do *AutoCorrect* (Passo a Passo *AutoCorrect*), disponível em:

https://drive.google.com/drive/folders/1xUYkB7EsXo-F9CXleLhZ0wrgK-SN3wtH?usp=drive_link. Todas as partes do site requerem que o usuário forneça o caminho para o arquivo a ser enviado. Após o upload, os arquivos são armazenados na nuvem por no máximo 24 horas. É importante que o nome dos arquivos esteja correto antes do envio (que sejam idênticos aos modelos disponíveis para Download e edição). Durante a geração dos gabaritos aleatorizados, o usuário deve indicar se deseja utilizar uma seed fixa. Em *Python*, uma seed fixa, especialmente em operações que envolvem números aleatórios, garante que os resultados sejam reproduzíveis. Em geral, funções que dependem de aleatoriedade, como geração de números ou amostragem de dados, produzem resultados diferentes a cada execução. No entanto, ao definir uma seed, o ponto de partida do gerador de números aleatórios é "travado", assegurando que os mesmos resultados sejam obtidos em execuções subsequentes (DOE 2024).

Na parte do código responsável pela geração dos gabaritos aleatorizados, o sistema usa o arquivo com os nomes dos alunos e o gabarito fornecido pelo professor. O *AutoCorrect* cria esse arquivo, chamado de "gabaritos aleatorizados", que é salvo automaticamente na pasta de downloads do usuário. Para personalizar as provas de cada aluno, o usuário deve abrir o arquivo em formato .xlsx, selecionar todas as planilhas e remover a primeira linha (que contém os nomes dos alunos). Após essa edição, o arquivo deve ser enviado novamente ao *AutoCorrect* para gerar as provas personalizadas.

RESULTADOS E DISCUSSÃO

Após a execução do código ou por meio do uso do *AutoCorrect*, são gerados os arquivos essenciais para atingir o objetivo final: correção das questões e obtenção de pontuações para cada estudante. O algoritmo utilizado é baseado na aleatorização controlada (seed), o que permite o embaralhamento de questões e gabaritos utilizando a biblioteca Random do *Python*, tornando o processo de ensino e aprendizagem mais dinâmico e eficiente (SANDVE et al. 2013). A Figura 10 exemplifica a geração de dois gabaritos aleatorizados para os alunos "Ana Clara" e "Pedro Henrique".

A	B	C	D	E	F	G	H	I	J	K
Ana Clara										
Questão	Questão C	Subitem 1	Subitem 2	Subitem 3	Subitem 4	Subitem 5	Subitem 6	Subitem 7	Subitem 8	Subitem 9
Q1	Q2	x	x		x			x		
Q2	Q1					5	0.2525			
Q3	Q3		1	2	3		5		7	

A	B	C	D	E	F	G	H	I	J	K
Pedro Henrique										
Questão	Questão C	Subitem 1	Subitem 2	Subitem 3	Subitem 4	Subitem 5	Subitem 6	Subitem 7	Subitem 8	Subitem 9
Q1	Q1					5	0.2525			
Q2	Q3		1	2	3		5		7	
Q3	Q2	x	x		x			x		

Figura 10. Gabaritos aleatorizados para determinados alunos da disciplina em questão.

Figure 10. Randomized answer sheets for specific students in the subject in question.

Um dos principais benefícios da aleatorização de gabaritos e questões é a possibilidade de criar versões únicas de provas para cada aluno, o que reduz

significativamente a chance de plágio ou colaborações indevidas durante os exames. Além disso, o uso de uma seed fixa garante que as versões geradas possam ser reproduzidas, permitindo que o professor tenha controle total sobre o processo, ao mesmo tempo que mantém a integridade da avaliação (DOE 2024).

Outro benefício é a capacidade de o professor utilizar um banco de questões extenso, permitindo que o sistema sorteie aleatoriamente questões diferentes para cada aluno. Isso personaliza as provas e oferece uma experiência de avaliação única para cada estudante, assegurando que todos sejam avaliados de forma justa e equitativa. Com isso, o professor pode criar uma avaliação diversificada sem a necessidade de gerar manualmente múltiplas versões da prova. Essa abordagem também permite o uso contínuo de um banco de questões ao longo de vários semestres, sem que os alunos recebam os mesmos testes repetidamente (KURDI et al. 2020).

Além de otimizar o tempo do docente, o uso de questões aleatorizadas incentiva o aprendizado, uma vez que cada aluno recebe desafios distintos, promovendo uma compreensão mais abrangente dos conteúdos (LIGUORI & WINKLER 2020). A combinação de gabaritos e questões aleatorizadas torna o ambiente de avaliação mais seguro e eficaz, destacando-se como uma prática de excelência no ensino superior, especialmente em cursos como Agronomia e Medicina Veterinária, que exigem avaliações robustas e diferenciadas.

Na geração de provas personalizadas para cada estudante, com base no gabarito aleatorizado, diversos algoritmos foram utilizados. Entre os principais, destacam-se: o uso de expressões regulares para identificar padrões e localizar as perguntas no documento; algoritmos para leitura e processamento de arquivos .csv e .xlsx; algoritmos de embaralhamento e mapeamento, que associam a ordem aleatorizada das questões ao gabarito de cada aluno; conversão de tabelas presentes nas questões em imagens, utilizando a biblioteca matplotlib; e a geração de documentos Word personalizados, inserindo texto e imagens conforme o gabarito específico de cada estudante. A Figura 11 ilustra um exemplo de prova, entregue ao aluno com questões aleatorizadas, onde a primeira questão corresponde à Q1, seguida de Q3 e Q2, exatamente conforme o gabarito aleatorizado gerado para esse aluno.

A correta correspondência entre a prova gerada e o gabarito é fundamental para garantir que a avaliação ocorra de maneira justa e precisa, assegurando que cada aluno seja avaliado com base no conjunto específico de questões que lhe foi atribuído. Este tipo de método de classificação automatizado e principalmente personalizável vem sendo usado por uma série de programadores ao redor do mundo, pelo fato de aplicar um padrão para classificar uma tarefa e gerenciar bibliotecas com o intuito de auxiliar na execução de testes em escolas (AKAHANE et al. 2015).

Após a aleatorização e geração das provas, ocorre a fase de realização do teste. Os alunos podem preencher o gabarito aleatorizado marcando o item correto com um "x" ou utilizando o caractere numérico correspondente, como 1, 2 ou 3 para cada questão. Essa automatização é essencial para evitar erros na leitura das respostas. Em questões que envolvem cálculo, o aluno deve inserir o valor calculado diretamente na questão, como por exemplo uma probabilidade (0,2525) associada a um evento.

Um aspecto importante é a opção antifraude, que penaliza respostas aleatórias. Se o aluno marca incorretamente uma alternativa, o sistema elimina uma resposta correta, assegurando que a avaliação reflete o conhecimento real e não a “sorte” (KASINATHAN et al. 2022).

Nome do Aluno: Pedro Henrique

Questão Q1:Q1. Os dados a seguir representam o rendimento de grãos de parcelas de soja, em gramas. 120; 100; 150; 250; 250; 155; 175; 185; 190; 200; 210; 215; 220. Pergunta-se: qual a escala de medida adotada? Qual o rendimento médio de grãos? A variação em g do conjunto? A conversão desta variação em percentual? Uma estimativa de erro na amostragem? Com base nisso, responda assinalando no gabarito apenas a(s) alternativa(s) correta(s): (1) Ordinal; 187; 45,74; 25% e 13. (2) Racional; 187; 45,74; 25% e 13; (3) Intervalar; 186,15; 45,74; 24,57% e 12,68; (4) Racional e Ordinal; 187; 45,90; 23,9% e 15; (5) Racional; 186,15; 45,74; 24,57% e 12,68; (6) Nenhuma das anteriores.

Questão Q2:Q3. Os dados a seguir representam a frequência de plantas de cada cultivar de soja: Zeus=10; Fibra= 9; Vênus=16; Valente= 18; Raio= 20 e Trovão =5. Com isso, assinale a/as alternativas corretas: (1) As cultivares fibra, raio e trovão apresentam a predominância de plantas; (2) Os extremos dessa distribuição são representados por Trovão e Raio; (3) O maior percentual de plantas ocorre em Raio; (4) O menor percentual de plantas é observado em Zeus; (5) Valente, Zeus e Fibra representam 47,42%; (6) Valente, Zeus e Fibra representam 47,44%; (7) A variável foi medida em escala nominal.

Figura 11. Exemplo de documento prova gerado para cada aluno apto a realizar a avaliação.

Figure 11. Example of a test document generated for each student eligible to take the assessment.

Diversos algoritmos foram implementados para este processo. O pd.read_excel é usado para ler as respostas dos alunos e os gabaritos, enquanto a função corrigir_prova compara as respostas com o gabarito e calcula a pontuação final. As notas são então organizadas e salvas em um arquivo Excel utilizando a função pd.DataFrame.to_excel da biblioteca *Pandas*, que estrutura os dados em linhas e colunas para facilitar o registro e análise dos resultados.

O uso de tecnologias digitais para a automação da avaliação dos resultados semestrais de uma disciplina torna-se importante pelo fato de reduzir ou eliminar a subjetividade na avaliação dos acadêmicos. O trabalho árduo na correção de uma grande quantidade de provas em um semestre, em um tempo hábil reduzido, pode condicionar professores ao erro de correção, prejudicando o desempenho de seus alunos (SANDVE et al. 2013, SMETANA & BELL 2012). Por outro lado, este sistema merece treinamento, já que de forma geral, os alunos das mais variadas instituições de ensino podem não estar adaptados com o método de avaliação, e podem ocorrer por exemplo erros de preenchimento no gabarito pelos estudantes (KASINATHAN et al. 2022, BLATTNER et al. 2023). A Figura 12, exemplifica como é atribuída a nota ao aluno, em comparação com o gabarito do professor.

A	B	C	D	E	F	G	H	I	J	K
QUESTAO	1	2	3	4	5	6	7	8	9	
Q1					5	0.2525				
Q2	x	x		x			x			
Q3	1	2	3		5		7			
Gabarito do professor										

A	B	C	D	E	F	G	H	I	J	K
Questão	Questão Original	Subitem 1	Subitem 2	Subitem 3	Subitem 4	Subitem 5	Subitem 6	Subitem 7	Subitem 8	Subitem 9
Q1	Q3		1	2	3		5		7	
Q2	Q2	x	x		x			x		
Q3	Q1						0.2121		7	

Aluna: Beatriz Santos

A	B	C	D	E	F	G	H	I	J	K
Questão	Questão Original	Subitem 1	Subitem 2	Subitem 3	Subitem 4	Subitem 5	Subitem 6	Subitem 7	Subitem 8	Subitem 9
Q1	Q2	x	x		x			x		
Q2	Q3					5			7	
Q3	Q1					4		0.1111		

Aluna: Mariana Souza

Figura 12. Exemplo de respostas dos alunos ao realizarem a avaliação.**Figure 12.** Example of student responses when completing the assessment.

As figuras demonstram como as notas foram atribuídas às alunas Beatriz e Mariana ao comparar suas respostas com o gabarito do professor. A aluna Beatriz acertou completamente a questão Q1 (que corresponde à questão 3 no gabarito original do professor), marcando corretamente os subitens 1, 2, 3, 5 e 7. Na questão Q2 (correspondente à Q2 no gabarito original), ela também acertou, marcando "x" nas colunas 1, 2, 4 e 7. No entanto, na questão Q3 (que corresponde à Q1 no gabarito do professor), Beatriz errou completamente, assinalando 0,2121 na coluna 6 e o número 7, resultando em zero pontos para essa questão. Com isso, Beatriz alcançou a nota final de 6,0. Já a aluna Mariana acertou completamente a questão Q1 (correspondente à Q2 no gabarito original), marcando corretamente "x" nas colunas 1, 2, 4 e 7. Na questão Q2 (que corresponde à Q3 no gabarito do professor), ela acertou parcialmente, marcando corretamente os subitens 5 e 7, mas não marcou os números 1, 2 e 3, o que resultou em uma pontuação parcial. Na questão Q3 (correspondente à Q1 no gabarito do professor), Mariana errou completamente, assinalando os números 4 e 0,1111, resultando em zero pontos para essa questão. Com isso, Mariana obteve a nota final de 4,6.

O sistema automatizado realiza essas comparações instantâneas, garantindo que as respostas incorretas anulem respostas corretas, assegurando uma avaliação justa. As respostas dos alunos são inseridas no arquivo "gabaritos_alunos", e atribui as notas de acordo com as regras estabelecidas. Esse método de correção não só evita possíveis fraudes ou tentativas de "chutar" respostas, mas também promove uma avaliação mais rigorosa, onde o conhecimento real do aluno é testado de forma eficaz. A anulação de respostas corretas em caso de marcação errada força os alunos a refletirem sobre suas respostas, reduzindo a probabilidade de marcações aleatórias e incentivando uma análise mais cuidadosa das alternativas (BASHITIALSHAAER et al. 2021, ELTAHIR et al. 2022).

Além disso, esse sistema automatizado de correção aumenta a transparência e a objetividade do processo avaliativo, pois elimina possíveis erros humanos ao conferir as provas manualmente. A padronização proporcionada por essa ferramenta é especialmente benéfica em avaliações de grande escala, como aquelas aplicadas em disciplinas com muitos alunos, garantindo que todos sejam avaliados de forma igualitária e consistente. Com a possibilidade de gerar diferentes versões de provas, o sistema também se torna mais robusto contra possíveis colaborações entre estudantes, uma vez que cada aluno pode receber uma combinação diferente de questões e gabaritos. Esse mecanismo também permite ao professor focar em aspectos mais estratégicos da avaliação, como a análise qualitativa do desempenho dos alunos, ao invés de se preocupar com a execução da correção propriamente dita. Ao reduzir o tempo e o esforço necessários para a correção, o professor pode dedicar mais atenção à revisão dos conceitos que se mostraram mais difíceis, promovendo um ciclo de feedback mais eficiente e alinhado às necessidades de aprendizado dos alunos (AYOUB-AL-SALIM & ALADWAN 2021)

Neste sistema, o professor digita as respostas dos alunos ou permite que os próprios alunos insiram as suas respostas. Para tanto basta o professor realizar o login no sistema *AutoCorrect* e clicar na opção “acessar gabarito”.

Ao clicar em acessar gabarito, o usuário se depara com a seguinte tela, sendo a senha o primeiro nome do aluno (Figura 13).

Para esta operação funcionar, o professor deve fazer o Upload do arquivo denominado de gabaritos_respostas. Este arquivo advém do arquivo gabaritos_aleatorizados e a partir disso, mediante um computador disponível, o próprio aluno pode inserir as suas respostas, de acordo com as questões presentes em sua avaliação.

A partir disso, o próprio aluno se encarrega de entrar com as respostas escolhidas para cada questão. Ademais da confecção de gabaritos aleatorizados, provas ajustadas com as questões ordenadas e correção automática das respostas, o *AutoCorrect* também permite a análise do desempenho dos estudantes por meio de cálculos de medidas descritivas e sua representação gráfica. Essas medidas descritivas são essenciais para entender a distribuição das notas e o comportamento geral dos alunos em relação à prova.

Questão	Subitem 1	Subitem 2	Subitem 3
Q1	None	None	None
Q2	None	None	None
Q3	None	None	None

Salvar Respostas

Figura 13. Acesso destinado ao aluno preencher as suas respostas da prova.

Figure 13. Access for students to fill in their test answers.

Na Figura 14, são apresentadas as estatísticas descritivas das provas, onde é possível observar a média geral das notas (7,33), o que indica um desempenho satisfatório na turma. A mediana, que é o valor central da distribuição, ficou ligeiramente abaixo da média, em 6,95, o que sugere uma leve inclinação para notas menores. O desvio padrão da média de 2,32 e a variância de 5,39 revelam que houve uma dispersão considerável nas notas dos alunos, o que pode indicar diferenças significativas no nível de conhecimento entre eles.

Além disso, o valor mínimo registrado foi 4,6, o que demonstra que nenhum aluno obteve uma nota extremamente baixa, enquanto o valor máximo foi 10,0, sugerindo que alguns alunos conseguiram atingir o desempenho máximo. A assimetria de 0,09 indica que as notas estão uniformemente distribuídas em torno da média da turma ($0,09 = 0$). Já a curtose negativa (-1,72) aponta que a distribuição de notas é platicúrtica, ou seja, mais achatada que a distribuição normal, indicando uma concentração maior de notas em torno da média. Essas medidas, ao serem aplicadas em conjunto com o sistema de correção automática, permitem uma análise precisa do desempenho dos alunos e oferecem ao professor insights valiosos sobre os pontos fortes e fracos da turma (WEISSGERBER et al. 2016). A combinação de gabaritos aleatorizados, correção automatizada e análise estatística torna o processo de avaliação mais robusto e orientado por dados, facilitando o planejamento de intervenções pedagógicas mais eficazes.

Estatísticas das Provas

- **Média Geral:** 7.33
- **Mediana:** 6.95
- **Desvio Padrão:** 2.3215512055520118
- **Variância:** 5.3896
- **Valor Mínimo:** 4.6
- **Valor Máximo:** 10.0
- **Curtose:** -1.7234863351468341
- **Assimetria:** 0.09262807507390926

Diferença da Nota em Relação à Média Geral

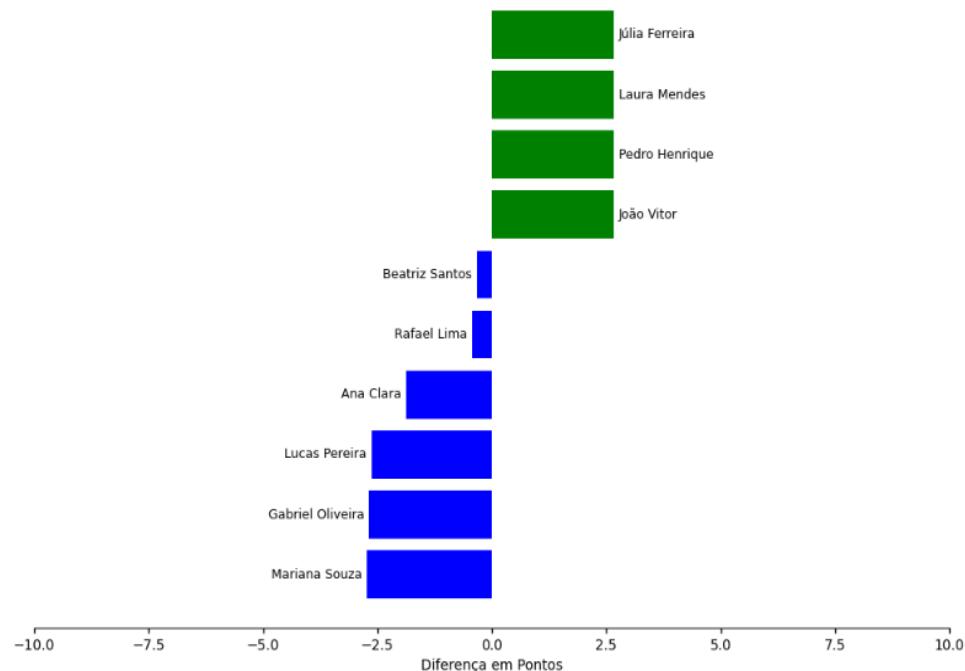


Figura 14. Cálculo das medidas descritivas e representação gráfica do desempenho dos estudantes.
Figure 14. Calculation of descriptive measures and graphical representation of student performance.

Além das estatísticas descritivas, o gráfico exemplifica a diferença da nota de cada aluno em relação à média geral da turma. Essa visualização permite identificar rapidamente quais alunos tiveram desempenho acima ou abaixo da média de 7,33. A partir da análise da figura, observamos que os alunos Júlia Ferreira, Laura Mendes, Pedro Henrique e João Vitor obtiveram notas superiores à média, com destaque para Júlia Ferreira, que teve o melhor desempenho, com uma diferença positiva de aproximadamente 2,5 pontos em relação à média. Por outro lado, alunos como Mariana Souza, Gabriel Oliveira, Lucas Pereira e Ana Clara tiveram desempenho abaixo da média, com Mariana Souza apresentando a maior diferença negativa, cerca de 3 pontos abaixo da média geral. Beatriz Santos e Rafael Lima obtiveram desempenhos próximos à média, indicando uma variação menor em relação aos outros alunos.

Esta representação gráfica é útil para o professor visualizar de forma clara o desempenho relativo dos alunos, permitindo identificar aqueles que podem necessitar de apoio adicional ou revisão de conteúdo. Ao mesmo tempo, alunos com desempenho acima da média podem ser reconhecidos por seu bom rendimento. A utilização de gráficos como este auxiliam na tomada de decisões pedagógicas direcionadas, facilitando o acompanhamento do progresso dos alunos de maneira visual e objetiva (SIAL 2021).

Por meio do desenvolvimento deste sistema se observa a sua utilidade na área da educação, promovendo ganhos precisos na avaliação dos estudantes. A Tabela 2 indica uma breve comparação entre os principais sistema utilizados no meio educacional no país e o *AutoCorrect*.

Tabela 2. Comparação entre os principais sistemas educacionais utilizados no meio educacional Brasileiro.

Table 2. Comparison between main educational systems used in Brazilian educational environment.

Sistema	Vantagens	Desvantagens
<i>Google Classroom</i>	Integração com ferramentas do google, gratuito para instituições educacionais, acesso via navegador.	Reduzida personalização, dependência de internet de boa qualidade.
<i>Microsoft Teams</i>	Integração com ferramentas da Microsoft, aulas síncronas e assíncronas, criar e corrigir provas via Forms.	Complexo para usuários recentes, versão gratuita tem limitações, consumo de recursos da máquina do usuário.
<i>Moodle</i>	Código aberto, ferramentas para criação de provas, não depende de servidores.	Conhecimento técnico capacitado para configuração, interface menos intuitiva, sistema lento.
<i>AutoCorrect</i>	Sistema de código aberto, flexível com as demandas do professor, eficiência e precisão na correção, utilização de arquivos simples, controle de plágio entre alunos.	Dependência de gabaritos rigorosos, necessidade de digitação das respostas dos alunos nos gabaritos.

Fonte: Adaptado de MUSTAFA & ALI 2023, RODILLAS et al. 2023.

Com base nessas análises e representações gráficas, fica evidente que a ciência de dados tem um impacto transformador na automação da correção de provas. A aplicação de algoritmos avançados para a aleatorização de gabaritos, correção automática e geração de estatísticas descritivas resultou em maior eficiência e precisão no processo avaliativo. Esse sistema automatizado não apenas minimiza o erro humano, como também promove uma avaliação justa, padronizada e transparente.

A visualização dos resultados em gráficos e tabelas facilita a identificação de padrões de desempenho dos alunos, permitindo aos professores realizarem intervenções pedagógicas eficazes (HAND 2019). Ao adotar essas ferramentas, instituições de ensino podem otimizar o tempo dos professores e garantir uma avaliação mais equilibrada e orientada por informações das turmas em cada semestre.

CONCLUSÃO

A automação na correção de provas utilizando ferramentas como *Python* e *AutoCorrect* online se mostrou uma solução robusta para gerenciar avaliações em grande escala. A integração de algoritmos e bibliotecas de análise de dados permitiu não só a correção eficiente e precisa, mas também ofereceu insights imediatos sobre o desempenho dos estudantes. Além disso, a personalização dos feedbacks e a eliminação de erros humanos tornaram o processo mais ágil e confiável. Esta modernização do processo de avaliação proporciona benefícios significativos tanto para professores, que podem focar em aspectos mais estratégicos do ensino, quanto para os alunos, que recebem uma avaliação justa e precisa em um menor tempo.

CONTRIBUIÇÕES DO AUTOR

Conceitualização, metodologia e análise formal, Paulo Henrique Cerutti; software e validação, Jefferson Luís Meirelles Coimbra; investigação, Mauro Bitencourt de Souza, Luan Tiago dos Santos Carbonari; recursos e curadoria de dados, Jefferson Luís Meirelles Coimbra; redação - preparação do rascunho original, Paulo Henrique Cerutti; redação - revisão e edição, Paulo Henrique Cerutti, Carlos Joquim Zacarias Júnior, Henrique de Sá Albino; visualização, Paulo Henrique Cerutti; supervisão e administração do projeto, Jefferson Luís Meirelles Coimbra; obtenção de financiamento, Jefferson Luís Meirelles Coimbra. Todos os autores leram e concordaram com a versão publicada do manuscrito.

FINANCIAMENTO

Este trabalho não foi apoiado por nenhuma agência de financiamento.

DECLARAÇÃO DO CONSELHO DE REVISÃO INSTITUCIONAL

Não aplicável a estudos que não envolvam humanos ou animais.

DECLARAÇÃO DE CONSENTIMENTO INFORMADO

Não aplicável porque este estudo não envolveu humanos.

DECLARAÇÃO DE DISPONIBILIDADE DE DADOS

Os dados podem ser disponibilizados mediante solicitação.

AGRADECIMENTOS

Agradecemos a Uviversidade do Estado de Santa Catarina (UDESC), ao Centro de Ciências Agroveterinárias (CAV), pelo apoio intelectual.

CONFLITOS DE INTERESSE

O presente manuscrito não possui conflitos de interesse.

REFERÊNCIAS

- ALDRIYE H et al. 2019. Automated grading systems for programming assignments: A literature review. *International Journal of Advanced Computer Science and Applications* 10: 250-259.
- AKAHANE Y et al. 2015. Design and evaluation of automated scoring: Java programming assignments. *International Journal of Software Innovation* 3: 18–32.
- AYOUB-AL-SALIM MI & ALADWAN K. 2021. The relationship between academic integrity of online university students and its effects on academic performance and learning quality. *Journal of Ethics in Entrepreneurship and Technology* 1: 43-60.
- BASHITALSHAAER R et al. 2021. Obstacle comparisons to achieving distance learning and applying electronic exams during COVID-19 pandemic. *Symmetry* 13: 99.
- BLATTNER A et al. 2023. One-Shot Grading: Design and Development of an Automatic Answer Sheet Checker. In: 2023 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE 562-566.
- DOE JO. 2024. Uso de Seed Fixa para Garantir Reprodutibilidade em Processos Computacionais. *Journal of Computational Reproducibility* 1: 1-5.
- ELTAHIR E et al. 2022. Implementation of E-exams during the COVID-19 pandemic: A quantitative study in higher education. *Plos One* 17: e0266940.
- HAND D. 2019. What is the purpose of statistical modeling? *Harvard Data Science Review* 1: 1-6. <https://doi.org/10.1162/99608f92.4a85af74>.
- HELMUS JJ & COLLIS SM. 2016. The *Python* arm radar toolkit (py-art), a library for working with weather radar data in the *Python* programming language. *Journal of Open Research Software* 4: 1-12.
- JAMES G et al. 2023. Introduction. In: *An Introduction to Statistical Learning*. Springer Texts in Statistics. Stanford: Department of Statistics. https://doi.org/10.1007/978-3-031-38747-0_1
- KASINATHAN V et al. 2022. ProctorEx: An Automated Online Exam Proctoring System. *Mathematical Statistician and Engineering Applications* 71: 876–889.
- KURDI G et al. 2020. A systematic review of automatic question generation for educational purposes. *International Journal of Artificial Intelligence in Education* 30: 121-204.
- LIGUORI E & WINKLER C. 2020. From offline to online: Challenges and opportunities for entrepreneurship education following the COVID-19 pandemic. *Entrepreneurship Education and Pedagogy* 3: 346-351.
- MASSARI CHAL et al. 2022. Veterinary anatomy during the COVID-19 pandemic in Brazil: Research focused on pedagogical practice. *International Journal of Morphology* 40:79-83.
- MUSTAFA AS & ALI N. 2023. The adoption and use of Moodle in online learning: A systematic review. *Information Sciences Letters* 12: 341-351.
- RODILLAS MJ et al. 2023. Mediating Effects of Teacher's Assessment Competencies on the Relationship Between the Use of the Zipgrade Application and Student's Achievement. *Psychology and Education: A Multidisciplinary Journal* 11: 862-876.
- ROLON-MÉRETTE D et al. 2016. Introduction to Anaconda and Python: Installation and setup. *Quantitative Methods for Psychology* 16: 3-11.

- SANDVE GK et al. 2013. Ten simple rules for reproducible computational research. *PLoS Computational Biology* 9: e1003285.
- SIAL M. 2021. A Brief Introduction to Regression Analysis and Its Types. *Asian Journal of Probability and Statistics* 13: 58-63.
- SMETANA LK & BELL RL. 2012. Computer simulations to support science instruction and learning: A critical review of the literature. *International Journal of Science Education* 34: 1337-1370.
- VAN DER AALST W. 2016. Data Science in Action. Berlin: Springer. https://doi.org/10.1007/978-3-662-49851-4_1.
- WEISSGERBER TL et al. 2016. From static to interactive: transforming data visualization to improve transparency. *PLoS Biology* 14: e1002484.
- ZHANG L et al. 2022. An automatic short-answer grading model for semi-open-ended questions. *Interactive Learning Environments* 30: 177-190.