

# Predição do Conceito ENADE dos Cursos de Computação no Brasil

Renan G. A. Sakashita<sup>1</sup>, Diego Bertolini<sup>1</sup>, André Luis Schwerz<sup>1</sup>

<sup>1</sup>Universidade Tecnológica Federal do Paraná (UTFPR)

Campo Mourão - Paraná - Brasil

**Abstract.** ENADE is a large-scale assessment applied to undergraduate students in Brazil. Their scores serve as a metric for assessing teaching quality, known as the ENADE Concept. A higher ENADE Concept rating can lead to increased visibility and investments in Higher Education Institutions. This work aims to evaluate the performance of Machine Learning techniques in predicting the ENADE Concept for computing courses in Brazil using microdata from the exam. Our experiments involve evaluating various classifiers with dataset variations and considering the impact of the temporal aspect of the exam on the prediction performance. We have identified that the Support Vector Machine produces the best prediction results. Additionally, our findings suggest that older training data leads to a reduction in prediction performance, indicating that course features may change over time.

**Resumo.** O ENADE é uma avaliação em larga escala aplicada a estudantes de graduação no Brasil. Suas pontuações servem como métrica para avaliar a qualidade do ensino, conhecida como Conceito ENADE. Uma classificação mais elevada do Conceito ENADE pode levar a maior visibilidade e investimentos nas Instituições de Ensino Superior. Este trabalho tem como objetivo avaliar o desempenho de técnicas de Aprendizado de Máquina na predição do Conceito ENADE para cursos de computação no Brasil utilizando microdados do exame. Os experimentos envolvem a avaliação de vários classificadores com variações no conjunto de dados e a consideração do impacto do aspecto temporal do exame no desempenho da predição. Identificou-se que o Support Vector Machine produz os melhores resultados de predição. Ademais, os resultados sugerem que dados de treinamento mais antigos reduzem o desempenho da predição, indicando que as características dos cursos mudam ao longo do tempo.

## 1. Introdução

O Exame Nacional de Desempenho dos Estudantes (ENADE) é uma prova realizada pelos concluintes de graduação no Brasil cujo o objetivo é avaliar seus cursos numa escala de 1 a 5, conhecida como Conceito ENADE (CE) [BRASIL 2023]. Esta avaliação é realizada em um ciclo avaliativo de três anos na qual, em cada ciclo, uma das três grandes áreas de conhecimento é avaliada. A prova apresenta tanto questões específicas, relacionadas à Diretriz Curricular Nacional do curso, quanto questões de escopo geral. Junto a ela, também é solicitado aos estudantes o preenchimento de um questionário socioeconômico.

Os dados coletados do exame, do questionário socioeconômico e do CE são disponibilizados de forma aberta em microdados pelo Instituto Nacional de Estudos e Pes-

quisas Educacionais Anísio Teixeira (INEP) [INEP 2023]. Dessa forma, é possível investigar o impacto das informações disponíveis dos cursos e de seus estudantes sobre o desempenho alcançado no exame. As descobertas podem ser úteis na elaboração de políticas de Educação Superior que fomentem o investimento de formação de recursos humanos [Marinho 1998].

Os resultados da prova do ENADE são utilizados em outros dois indicadores de qualidade: no Conceito Preliminar do Curso (CPC) e no Indicador de Diferença entre os Desempenhos Observado e Esperado (IDD). O CPC atribui ao CE e ao IDD, em sua fórmula, os pesos de 20% e 35%, respectivamente [Inep 2023]. Desse modo, destaca-se a relevância da prova e a importância de obter-se um alto CE, visto sua participação em outro importante indicador.

Há diversos trabalhos que investigam os microdados do ENADE, como descrito em [Barbosa et al. 2023]. A grande maioria deles trata de forma individualizada o desempenho dos estudantes no exame, enquanto o CE ainda é pouco explorado. Além disso, no melhor do conhecimento, nenhum estudo investigou o impacto da evolução temporal das características dos cursos no resultado da predição.

Os objetivos deste trabalho são: propor uma abordagem baseada em Aprendizado de Máquina para avaliar o impacto das características coletadas nos microdados do ENADE para a predição do Conceito ENADE; e determinar o impacto das características dependentes do tempo na predição. Para os experimentos, são utilizados os microdados do ENADE das edições de 2011, 2014, 2017 e 2021 dos cursos na área de computação. Um extenso pré-processamento que inclui seleção, limpeza, integração e transformação foi realizado para extração de características dos cursos e criação do conjunto de dados usado nos experimentos de predição.

Na Seção 2 serão apresentados os trabalhos relacionados. Na Seção 3, será exposto o método utilizado para o processamento e avaliação dos microdados. Por fim, na Seção 4 mostra-se os resultados e as discussões dos experimentos e na Seção 5 sintetiza-se as conclusões deste trabalho.

## **2. Trabalhos Relacionados**

Um amplo mapeamento sistemático sobre os trabalhos e as técnicas usadas na análise dos dados do ENADE foi apresentado em [Barbosa et al. 2023]. Eles identificaram na literatura as diversas técnicas estatísticas e de aprendizado de máquina que foram utilizadas com os microdados do ENADE, em especial, para a predição do CE.

O Aprendizado Supervisionado foi usado em [Nunes et al. 2018, Martins et al. 2021] para predição do CE. Nunes et al. avaliaram cursos de Engenharia de Produção nas edições 2011 e 2014, aplicando o MLP para analisar a influência e a capacidade das notas de formação geral e de componente específico de afetar e prever o CE dos cursos. Como resultado, alcançaram uma acurácia de 99,997% na predição do CE [Nunes et al. 2018]. Martins et al. usaram DT para prever o CE em uma pequena quantidade de cursos do estado de Goiás analisando dados da edição 2012 [Martins et al. 2021]. Como resultado, obtiveram uma acurácia de 96%. Embora relacionados, ambos trabalhos não podem ser comparados diretamente com a presente pesquisa porque simplificam o problema ao usar as notas parciais dos estudantes para prever o CE.

Há também pesquisas que utilizaram Aprendizado Não Supervisionado sobre os microdados do ENADE para observar fatores determinantes para o CE. Em [Silva et al. 2021], analisou-se, a partir do algoritmo K-Means, a relação da categoria administrativa e da região dos cursos de Medicina com seus respectivos CE nas edições 2010, 2013 e 2016. Em [Figueiro et al. 2018], foram utilizados os dados dos cursos do Consórcio das Universidades Comunitárias Gaúchas (COMUNG) na edição 2014 e o algoritmo de agrupamento hierárquico para análise da relação entre o número de inscritos, o número de participantes, as notas de formação geral e de componente específico com seus respectivos CE.

A grande maioria dos esforços de pesquisa têm observado o desempenho dos estudantes no exame. Rosa et al. utilizaram os dados dos cursos de computação, das edições 2008 a 2017, treinando uma série de classificadores para predição do desempenho dos participantes [Rosa et al. 2021]. Capelari estendeu esse trabalho avaliando novas classes e classificadores [Capelari 2022]. Em [Vieira et al. 2022], realizou-se uma análise evolutiva dos grupos obtidos pelo K-Means ao longo das edições do exame, observando aspectos relevantes para o desempenho dos estudantes dos cursos de computação. A análise descritiva também foi utilizada com os dados dos cursos de medicina em [Silva et al. 2021]. Em [Capelari and Schwerz 2021], ela foi usada para mostrar o perfil socioeconômico dos estudantes dos cursos de computação da região sul do Brasil.

No melhor do conhecimento, nenhuma contribuição anterior investiga a predição do CE com uma grande quantidade de características extraídas dos microdados e o impacto da evolução temporal das características dos estudantes que compõem os cursos de computação na predição do CE.

### **3. Método de Pesquisa**

Neste trabalho utilizou-se o método *Knowledge Discovery in Databases* (KDD) [Frawley et al. 1992] para guiar o estudo quantitativo dos microdados do ENADE. O KDD visa orientar o processo de extração de informações não-triviais, desconhecidas e potencialmente relevantes em banco de dados. Desta forma, ele inclui etapas para o tratamento de informações e aplicação de técnicas com diferentes objetivos, como classificação, agrupamento e reconhecimento de padrões. As etapas do KDD são: pré-processamento, mineração, e avaliação.

#### **3.1. Pré-processamento**

O pré-processamento inclui a seleção, limpeza, transformação e integração de dados. Cada uma destas etapas é descrita em detalhes nas seções seguintes. Como resultado, gerou-se um conjunto de dados com características dos cursos de graduação da área de computação no Brasil, processadas a partir dos microdados do ENADE.

##### **3.1.1. Seleção de Dados**

A etapa de seleção levou em consideração a reformulação dos microdados do ENADE, disponível a partir de 2022, devido à Lei Geral de Proteção de Dados (LGPD) [INEP 2023]. Por causa disso, as respostas das questões do questionário socioeconômico e outras informações foram aleatorizadas entre si e distribuídas por diferentes

arquivos de forma que não é possível identificar uma série de informações de um aluno, mas sim informações gerais de um curso de graduação<sup>1</sup>. Por exemplo, a partir dos microdados não pode-se afirmar que um determinado aluno de cor/raça branca tenha 19 anos, mas pode-se inferir que 60% dos alunos de um curso têm cor/raça branca e que a idade média é de 19 anos.

Originalmente, para cada edição do exame, o conjunto de dados é dividido em arquivos que contém o código do curso acompanhado por alguma informação específica como, por exemplo, as respostas do questionário socioeconômico ou a cor/raça declaradas pelos participantes. Por causa dessa fragmentação, foi preciso ordenar o conteúdo de cada arquivo pelo código do curso, para junção e, então, formação do conjunto de dados usado neste trabalho.

Os questionários socioeconômicos não estão padronizados ao longo dos anos. Os dados das edições 2005 e 2008 apresentam inúmeras diferenças com as edições mais recentes e, por esse motivo, optou-se por não usá-los neste trabalho. Para as demais edições, selecionou-se um subgrupo de atributos comuns em todas as edições (2011, 2014, 2017 e 2021). No total obteve-se 36 atributos: 6 pertencentes aos cursos, 24 do questionário socioeconômico e 6 derivados<sup>2</sup>, apresentados na Tabela 1.

**Tabela 1. Atributos selecionados**

Categoria	Atributos
Curso	Categoria da Organização Acadêmica, Turno, Área, UF, Região, Categoria Administrativa.
Socioeconômico	UF do ensino médio, Situação financeira, Condição das salas práticas, Cor ou Raça, Situação de trabalho, Formação, Estado civil, Bolsa de estudos, Tipo da escola do ensino médio, Moradia, Disponibilidade dos professores, Modalidade do ensino médio, Quantidade de pessoas morando junto, Nível de escolaridade do Pai, Domínio dos professores, Idade, Política de ingresso, Nível de escolaridade da Mãe, Condição das Salas, Sexo, Quantidade de livros, Renda, Horas de estudo, Plano de ensino.
Derivados	Quantidade de notas inválidas, de alunos faltantes, de participantes, Conclusão do ensino médio, Anos de preparação, Anos de graduação.

O enquadramento é o processo pelo qual cada curso é vinculado à sua respectiva área de avaliação no ENADE [Inep 2022b]. Neste trabalho, utilizou-se as áreas de enquadramento relacionadas à computação, exibidas na Tabela 2.

### 3.1.2. Limpeza de Dados

Os conjuntos de dados das edições foram analisados em relação à existência de valores não informados a fim de mensurar seu impacto na qualidade dos dados. No entanto, como

<sup>1</sup>Esse processo é descrito no arquivo Manual do Usuário presente nos microdados [INEP 2023].

<sup>2</sup>Esses atributos foram formados a partir de atributos presentes nos microdados. Por exemplo, “Conclusão do ensino médio” é a diferença entre o ano da edição e o ano em que o estudante concluiu o ensino médio.

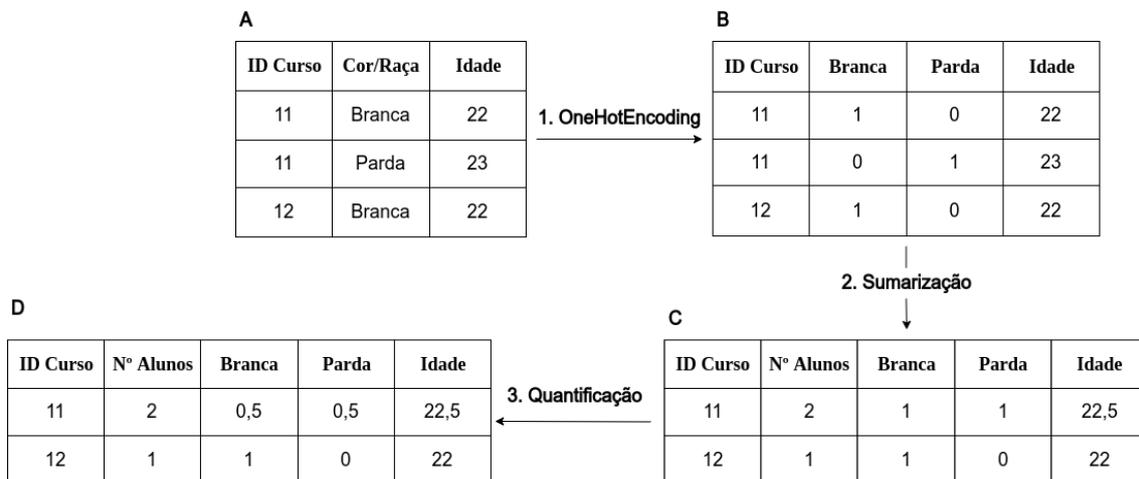
**Tabela 2. Áreas de enquadramento relacionadas à computação**

Sigla	Grau	Área de enquadramento
ADS	Tecnologia	Análise e Desenvolvimento de Sistemas
BCC	Bacharelado	Ciência da Computação
EC	Bacharelado	Engenharia da Computação
GTI	Tecnologia	Gestão da Tecnologia da Informação
LCC	Licenciatura	Ciência da Computação
RC	Tecnologia	Redes de Computadores
SI	Bacharelado	Sistemas de Informação

citou-se anteriormente, as informações foram aleatorizadas entre si, implicando que a exclusão de uma instância por causa de um valor defeituoso levaria à provável eliminação de valores de diferentes alunos de um curso. Desta forma, valores nulos foram tratados como valores não informados.

### 3.1.3. Transformação de Dados

A granularidade original dos microdados do ENADE refere-se às informações dos estudantes. Neste estudo, para obtenção de instâncias referentes aos cursos, houve a combinação da técnica de *one-hot encoding* com procedimentos de sumarização e de quantificação ilustrados na Figura 1. Por exemplo, dado uma amostra **A** extraída dos microdados, obtém-se o conjunto **B** por meio do *one-hot encoding* de variáveis categóricas; em seguida, o conjunto **C** é obtido por meio da soma dos valores dos atributos nominais e da média dos numéricos; por fim, obtém-se um conjunto **D** por meio da divisão das somas dos valores categóricos pelo número de alunos efetivos do curso. Após estas operações de transformação, fez-se também a normalização dos dados utilizando o método Z-Score [Rout 2023].



**Figura 1. Processo ilustrativo de transformação de dados de alunos em cursos**

Inicialmente, o conjunto possuía 205.837 instâncias considerando todos os estudantes participantes das edições 2011, 2014, 2017 e 2021. Após o processo de transformação descrito acima, obteve-se 5.838 instâncias representando cada curso presente no conjunto. A Tabela 3 sintetiza as informações das edições quanto ao número de estudantes e cursos, e ao número de cursos por área de enquadramento<sup>3</sup>.

**Tabela 3. Dimensão do conjunto de dados**

Edição	#estudantes	#cursos	ADS	BCC	EC	GTI	LCC	RC	SI
2011	41.023	1.222	259	386	14	0	39	167	357
2014	51.774	1.505	352	304	119	0	54	197	479
2017	51.009	1.605	367	308	145	118	55	158	454
2021	62.031	1.506	482	319	0	147	56	114	388
Total	205.837	5.838	1.460	1.317	278	265	204	636	1.678

### 3.1.4. Integração de Dados

Originalmente, os microdados não contêm o CE dos cursos. Por este motivo, após a etapa de transformação, tornou-se necessária a integração dos conjuntos de dados disponibilizados pelo ENADE que informam o CE [Inep 2020]. Entretanto, não é possível a integração dos conceitos dos cursos de 2011 e de 2014 porque, nos dados destas edições, o código do curso é omitido. Para resolver esse problema, o CE dos cursos de todas as edições foram calculados usando a Equação 1 [Inep 2022a].

$$NC_c = 0,25 \cdot NP_{FG_c} + 0,75 \cdot NP_{CE_c}^4, \quad (1)$$

onde  $NC_c$ ,  $NP_{FG_c}$  e  $NP_{CE_c}$  correspondem, respectivamente, à nota dos concluintes, nota padronizada em formação geral e nota padronizada em componente específico do curso de graduação  $c$ . Como uma forma de validar a implementação, os resultados obtidos na Equação 1 foram contrapostos com os dados das edições 2017 e 2021.

A proporção de cursos com CE 1, 2, 3, 4 e 5 pelas edições 2011, 2014, 2017 e 2021 é apresentada na Tabela 4. Percebe-se um desbalanceamento entre as classes. Há, por exemplo, uma predominância do CE 3 representando por volta de 40% dos cursos nas quatro edições enquanto os CE 1 e CE 5 possuem um pequeno número de instâncias.

### 3.2. Mineração de Dados

Para avaliar a predição do CE, os seguintes classificadores foram utilizados: Árvore de Decisão (DT, do inglês *Decision Tree*), k-Vizinhos Mais Próximos (k-NN, do inglês *k-Nearest Neighbors*), Floresta Aleatória (RF, do inglês *Random Forest*), Máquina de Vetores de Suporte (SVM, do inglês *Support Vector Machine*) e Perceptron Multicamadas

<sup>3</sup>A área GTI não está presente em 2011 e 2014, pois foi criada em 2017, e a área EC foi movida para outro ciclo avaliativo em 2021, não estando presente na edição.

<sup>4</sup>Obtém-se dessa equação um valor contínuo que é substituído por um número de 1 a 5 de acordo com intervalos definidos em [Inep 2022a], constituindo o CE do curso.

**Tabela 4. Distribuição do CE ao longo das edições**

Edição	CE 1	CE 2	CE 3	CE 4	CE 5
2011	58 (4,7%)	317 (25,9%)	549 (44,9%)	201 (16,4%)	97 (7,9%)
2014	69 (4,6%)	469 (31,2%)	593 (39,4%)	300 (19,9%)	74 (4,9%)
2017	60 (3,7%)	465 (29%)	636 (39,6%)	349 (21,7%)	95 (5,9%)
2021	64 (4,2%)	502 (33,3%)	604 (40,1%)	281 (18,7%)	55 (3,7%)

(MLP, do inglês *Multilayer Perceptron*). Estes classificadores são bem conhecidos na área de Aprendizado de Máquina e usados para uma grande variedade de problemas. Além disso, todos eles já foram usados em trabalhos relacionados com os microdados do ENADE de acordo com [Barbosa et al. 2023].

Para o desenvolvimento dos modelos, foram utilizadas a linguagem de programação Python e a biblioteca scikit-learn [Pedregosa et al. 2011]. Para manter a reprodutibilidade da pesquisa, o código-fonte e os dados dos experimentos podem ser encontrados a partir do repositório<sup>5</sup>.

### 3.3. Avaliação

Durante o treinamento e avaliação dos modelos foram aplicadas as seguintes técnicas: HalvingGridSearchCV, F1-Score e validação cruzada. Nota-se que, antes da avaliação do conjunto, os dados de treino e teste foram normalizados de forma separada usando o Z-Score.

O HalvingGridSearchCV [Scikit-Learn 2023] é um algoritmo capaz de estimar a melhor combinação de parâmetros para um modelo de Aprendizado de Máquina, considerando uma base de dados e um conjunto de hiperparâmetros previamente fornecidos. Neste trabalho, foram usados os hiperparâmetros exibidos na Tabela 5.

**Tabela 5. Hiperparâmetros dos classificadores**

Modelo	Hiperparâmetros
DT	<b>criterion:</b> [gini, entropy], <b>min_samples_leaf:</b> [10, 30, 50], <b>ccp_alpha:</b> [0.001, 0.0015, 0.0017, 0.002]
k-NN	<b>n_neighbors:</b> [3, 5, 10], <b>weights:</b> [uniform, distance], <b>metric:</b> [euclidean, manhattan, nan_euclidean]
RF	<b>n_estimators:</b> [60, 100, 200], <b>criterion:</b> [gini, entropy], <b>min_samples_leaf:</b> [10, 30, 50], <b>ccp_alpha:</b> [0.001, 0.0015, 0.0017, 0.002]
SVM	<b>C:</b> [0.01, 0.1, 1, 10, 100], <b>kernel:</b> [linear, poly, rbf], <b>gamma:</b> [0.1, 1, 10], <b>tol:</b> [0.1, 0.001, 0.00001]
MLP	<b>hidden_layer_sizes:</b> [10, 20, 10], <b>activation:</b> [relu, tanh], <b>solver:</b> [adam, sgd], <b>learning_rate_init:</b> [0.001, 0.01, 0.1], <b>tol:</b> [0.1, 0.001, 0.0001], <b>alpha:</b> [0.0001, 0.01, 0.1], <b>batch_size:</b> [32, 60, 100]

<sup>5</sup><https://github.com/RenanGAS/ML-ENADE>

Para avaliar o desempenho dos classificadores, adotou-se o F1-Score que é uma métrica de avaliação adequada para conjuntos de dados desbalanceados. Ela considera o desempenho médio das predições de cada classe do problema, ao invés da porcentagem de acertos geral como é feito na acurácia.

Nos experimentos preliminares, adotou-se a validação cruzada ( $n = 5$ ) para avaliar a generalização dos modelos abordados. Esta técnica propõe a divisão de um conjunto de dados em  $n$  partes de tamanhos semelhantes, de forma que ocorram  $n$  ciclos de avaliação do desempenho do conjunto, sendo que em cada ciclo é utilizado uma parte diferente para teste e o restante para treinamento do modelo em questão.

## 4. Experimentos e Discussões

Para análise dos microdados do ENADE, dois conjuntos de experimentos foram elaborados. O primeiro conjunto, descrito na Seção 4.1, tem como objetivo consolidar uma base de referência na aplicação dos algoritmos de predição sobre o conjunto de dados e suas divisões. O segundo conjunto de experimentos, apresentado na Seção 4.2, tem foco em avaliar o impacto das características dependentes do tempo na predição.

### 4.1. Experimentos Preliminares

Foram realizados três experimentos usando validação cruzada com cinco partições para todos os classificadores destacados na Seção 3.2. Para cada um deles, registrou-se a média e o desvio padrão do F1-Score das iterações em cada modelo de classificação.

Inicialmente, utilizou-se o conjunto de dados com as edições 2011, 2014, 2017 e 2021, sem quaisquer alterações, obtendo os resultados exibidos na Tabela 6. As melhores taxas foram atingidas pelos modelos SVM e MLP, com os valores de 46,84% e 46,55%, respectivamente.

**Tabela 6. Comparativo das médias do F1-Score obtidas usando cinco classificadores para o conjunto total de dados**

DT	k-NN	RF	SVM	MLP
38,17 ± 1,14	43,47 ± 2,06	43,11 ± 2,30	<b>46,84 ± 2,84</b>	<b>46,55 ± 2,00</b>

Em seguida, elaborou-se dois experimentos considerando a divisão do conjunto de dados por edição e por área de enquadramento, com o objetivo de melhorar o desempenho da predição.

No primeiro, formou-se conjuntos para cada edição, treinando e testando os modelos com dados do mesmo ano. Por uma questão de clareza, a Tabela 7 sintetiza apenas os melhores resultados que foram obtidos com o classificador SVM. Todos os outros resultados podem ser obtidos no repositório. Destaca-se o desempenho da edição 2021 com 51,32% de média e baixo desvio padrão, enquanto as outras edições se mantiveram numa faixa inferior e com taxas semelhantes entre si considerando os desvios padrões. Uma das possíveis razões de sua distinção é que, na edição 2021, há o maior número de estudantes participantes, mesmo com a saída da área de enquadramento da EC e, conseqüentemente, a redução do número de cursos avaliados, como mostra a Tabela 3.

**Tabela 7. Comparativo das médias do F1-Score obtidas usando o classificador SVM para o conjunto de dados dividido pelas edições do exame**

2011	2014	2017	2021
46,23 ± 4,50	45,96 ± 3,28	43,52 ± 1,12	<b>51,32 ± 1,82</b>

No segundo experimento, o conjunto foi dividido em áreas de enquadramento. Novamente, os melhores resultados foram obtidos pelo classificador SVM, sendo apresentados na Tabela 8. Destaca-se os resultados das áreas de ADS, BCC e SI, que são, respectivamente, 47,84%, 50,49% e 48,65%. Mais uma vez, a diferença entre as médias observadas pode estar relacionada ao número de instâncias avaliadas. A Tabela 3 mostra que as áreas de enquadramento ADS, BCC e SI são aquelas que possuem o maior número de cursos.

**Tabela 8. Comparativo das médias do F1-Score obtidas usando o classificador SVM para o conjunto de dados dividido pelas áreas de enquadramento**

ADS	BCC	EC	GTI	LCC	RC	SI
<b>47,84 ± 2,14</b>	<b>50,49 ± 1,72</b>	45,79 ± 5,51	35,43 ± 6,15	44,97 ± 7,18	42,59 ± 2,88	<b>48,65 ± 2,74</b>

Com relação aos estudos correlatos, torna-se inviável realizar uma comparação apropriada, uma vez que as pesquisas apresentadas por [Martins et al. 2021] e [Nunes et al. 2018] utilizam notas parciais dos estudantes como características de treinamento dos modelos e, conseqüentemente, uma alta acurácia nos testes.

#### 4.2. Aspecto Temporal do Conjunto

Os dados usados neste trabalho são um recorte dos cursos da área da computação produzidos ao longo de dez anos. A hipótese desta pesquisa é que há uma evolução das características dos cursos que inviabilizaria um modelo de predição único. Esse fenômeno é conhecido na literatura como *concept drift* [Webb et al. 2016]. Para avaliar essa hipótese, uma série de experimentos foram realizados levando em consideração o aspecto temporal do conjunto de dados. Diferentemente dos experimentos preliminares, os modelos foram treinados usando todas as instâncias de um ou mais anos e, para os testes, utilizou-se dados de um ano posterior. Por exemplo, no primeiro experimento, utilizou-se as instâncias de 2011 para o treinamento e todas as instâncias de 2014 para o teste. Observa-se que, com os conjuntos fixos de treinamento e teste, a validação cruzada e o cálculo do desvio padrão não são necessários.

Os resultados dos experimentos do *concept drift* são apresentados na Tabela 9. Com o conjunto original, em geral, o desempenho é mais promissor quando a edição 2021 é designada para teste e as demais para treino, alcançando as taxas de 48,19%, 52,28% e 48,06%, sendo superior em comparação aos resultados da Tabela 6, com a maior média de 46,84% obtida usando o SVM. Para os demais resultados, em que as edições 2014 e 2017 são avaliadas, não houve uma melhoria no desempenho. Apesar dos valores da métrica F1-Score permanecerem na mesma faixa dos experimentos com o conjunto inteiro, observa-se que a edição 2011, quando usada para treinamento, reduz o

desempenho do modelo. A Tabela 9 também exibe os resultados de experimentos em que o conjunto de dados foi dividido em áreas de enquadramento. Optou-se por exibir apenas as áreas de enquadramento com maior número de instâncias (ADS, BCC e SI). Com a exceção do caso na área BCC em que houve o treinamento com os dados de 2011 e teste com os de 2014, alcançando uma taxa de 49,84%, todos os demais resultados são similares aos obtidos com o conjunto original. Observa-se também que o SVM não apresenta o melhor desempenho em todos os cenários, sendo superado por outros classificadores (k-NN, MLP, RF).

**Tabela 9. Comparativo do F1-Score obtido considerando o aspecto temporal dos dados**

Treino	Teste	Conjunto original	ADS	BCC	SI
2011	2014	39,98 (k-NN)	29,87 (MLP)	<b>49,84 (k-NN)</b>	41,33 (k-NN)
2014	2017	41,82 (SVM)	38,10 (MLP)	42,37 (SVM)	37,27 (SVM)
2017	2021	<b>48,19 (SVM)</b>	49,70 (MLP)	51,99 (RF)	52,70 (SVM)
2011, 2014	2017	42,74 (SVM)	43,69 (SVM)	45,98 (SVM)	40,33 (MLP)
2014, 2017	2021	<b>52,28 (SVM)</b>	50,75 (MLP)	54,95 (RF)	47,35 (SVM)
2011, 2014, 2017	2021	<b>48,06 (MLP)</b>	47,12 (SVM)	55,59 (SVM)	48,33 (SVM)

## 5. Conclusões

Neste trabalho apresentou-se um estudo abrangente com várias características extraídas dos microdados do ENADE com o objetivo de prever o CE dos cursos e avaliar o impacto do aspecto temporal na predição. Os resultados apontam que o desempenho dos modelos varia de 38,17% a 46,84% e os experimentos usando as divisões (edição e área de enquadramento) do conjunto de dados sugerem que o desempenho pode ser melhorado quando há maior número de instâncias. Em relação ao aspecto temporal, identificou-se que os dados da edição 2011 reduzem o desempenho dos modelos de predição sugerindo que o conjunto de dados possui um *concept drift*. Além disso, para realização dos experimentos, foi produzido um conjunto de dados com características de cursos da área de computação que pode ser usado pela comunidade para realização de outras investigações.

Como trabalhos futuros, espera-se melhorar o desempenho da predição com o enriquecimento dos dados com outros conjuntos de dados públicos e a combinação dos classificadores. Os microdados do Censo da Educação Superior (Censup) e do IDD seriam interessantes, visto que os dados do Censup são utilizados como parâmetro no cálculo de outro indicador de qualidade, o Índice Geral de Cursos Avaliados da Instituição (IGC); e os microdados do IDD devido à pesquisa feita em [Filho 2018] na qual o conjunto mostrou-se eficaz na predição do desempenho dos cursos de Engenharia da Produção no ENADE.

## Agradecimentos

Os autores agradecem a Universidade Tecnológica Federal do Paraná – Campus Campo Mourão (Editais PROPPG nº 02/2022 e DIRPPG-CM nº 01/2023) pelo apoio financeiro.

## Referências

- [Barbosa et al. 2023] Barbosa, P., Damazio, G., Carvalho, W., and Carmo, R. (2023). O sucesso não é apenas uma questão de sorte: um mapeamento sistemático sobre técnicas de análise do enade da área de computação. In *Anais do III Simpósio Brasileiro de Educação em Computação*, pages 59–68, Porto Alegre, RS, Brasil. SBC.
- [BRASIL 2023] BRASIL (2023). Instituto nacional de estudos e pesquisas educacionais anísio teixeira. <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enade> [Acessado em 01-jul-2023].
- [Capelari 2022] Capelari, L. O. O. (2022). Predição do desempenho no enade dos discentes de computação. Universidade Tecnológica Federal do Paraná.
- [Capelari and Schwerz 2021] Capelari, L. O. O. and Schwerz, A. L. (2021). O perfil socioeconômico dos concluintes de computação do sul do brasil. *Anais do Computer on the Beach*, 12:133–140.
- [Figueiro et al. 2018] Figueiro, M., Vista, N., Barasuol, J., Chicon, P., and Ansuji, A. (2018). Análise de agrupamento hierárquico aplicada aos microdados do enade do curso de graduação em ciência da computação. *Revista Eletrônica Argentina-Brasil de Tecnologias da Informação e da Comunicação*, 1(8).
- [Filho 2018] Filho, R. L. A. (2018). Análise do desempenho acadêmico enem/enade como instrumento de qualificação dos cursos de engenharia de produção no brasil. Universidade Tecnológica Federal do Paraná.
- [Frawley et al. 1992] Frawley, W. J., Piatetsky-Shapiro, G., and Matheus, C. J. (1992). Knowledge discovery in databases: An overview. *AI Magazine*, 13(3):57.
- [Inep 2020] Inep (2020). Indicadores de qualidade da educação superior. <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/indicadores-educacionais/indicadores-de-qualidade-da-educacao-superior> [Acessado em 08-jul-2023].
- [Inep 2022a] Inep (2022a). Nota técnica nº 7/2022/cgcqes/daes. [https://download.inep.gov.br/educacao\\_superior/enade/notas\\_tecnicas/2019/nota\\_tecnica\\_n\\_7\\_2022\\_CGCQES\\_DAES\\_metodologia\\_calculo\\_conceito\\_enade\\_2021.pdf](https://download.inep.gov.br/educacao_superior/enade/notas_tecnicas/2019/nota_tecnica_n_7_2022_CGCQES_DAES_metodologia_calculo_conceito_enade_2021.pdf) [Acessado em 08-jul-2023].
- [INEP 2023] INEP (2023). Microdados do enade. <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enade> [Acessado em 01-jul-2023].
- [Inep 2023] Inep (2023). Nota técnica nº 4/2023/cei/cggi/daes. [https://download.inep.gov.br/educacao\\_superior/enade/notas\\_tecnicas/2023/nota\\_tecnica\\_n\\_4\\_2023\\_cei\\_cggi\\_daes\\_descricao\\_da\\_metodologia\\_de\\_calculo\\_do\\_cpc.pdf](https://download.inep.gov.br/educacao_superior/enade/notas_tecnicas/2023/nota_tecnica_n_4_2023_cei_cggi_daes_descricao_da_metodologia_de_calculo_do_cpc.pdf) [Acessado em 27-ago-2023].
- [Inep 2022b] Inep, A. (2022b). Verificação do enquadramento automático vai até 31/8. <https://www.gov.br/inep/pt-br/assuntos/noticias/enade/verificacao-do-enquadramento-automatico-vai-ate-31-8> [Acessado em 08-jul-2023].

- [Marinho 1998] Marinho, A. (1998). O aporte de recursos públicos para as instituições federais de ensino superior. *Revista de Administração Pública*, 32(4):83 a 93.
- [Martins et al. 2021] Martins, E. R., Alarcão, D. T. A., and Vieira, S. L. (2021). Descoberta de conhecimento aplicado ao enade. In Martins, E. R., editor, *Engenharia de Produção: Gestão de Qualidade, Produção e Operações - Vol. 2*, pages 77–90. Editora Científica Digital.
- [Nunes et al. 2018] Nunes, R. T. M., Quadrelli, G., and Nazaré, T. B. (2018). Cálculo preditivo de classificação das notas do enade utilizando redes neurais artificiais. *Revista de Tecnologia Aplicada*, 7(2):71–81.
- [Pedregosa et al. 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- [Rosa et al. 2021] Rosa, E. R., Ferreira, D. J., da Silva, N. F. F., and Assis, A. (2021). Estudo exploratório através de análises longitudinais aplicado à ciência da computação a partir da base de dados do enade. *Revista Brasileira de Informática na Educação*, 29:1463–1486.
- [Rout 2023] Rout, A. R. (2023). Z-score in statistics. <https://www.geeksforgeeks.org/z-score-in-statistics/> [Acessado em 05-jul-2023].
- [Scikit-Learn 2023] Scikit-Learn (2023). 3.2. tuning the hyper-parameters of an estimator. [https://scikit-learn.org/stable/modules/grid\\_search.html#searching-for-optimal-parameters-with-successive-halving](https://scikit-learn.org/stable/modules/grid_search.html#searching-for-optimal-parameters-with-successive-halving) [Acessado em 09-jul-2023].
- [Silva et al. 2021] Silva, J. C. P., de Oliveira, E. J., Neto, F. M. M., Santos, A. M., e Silva, L. C., Marques, L. T., Mendes, I. M., de Sousa Neto, A. F., Silva, J. C. P., and Silva, S. D. (2021). Data analysis of the performance of brazilian higher education medicine courses. In *Proceedings of the 10th Euro-American Conference on Telematics and Information Systems*, EATIS '20, New York, NY, USA. ACM.
- [Vieira et al. 2022] Vieira, A., Bertolini, D., and Schwerz, A. (2022). Análise do desempenho no enade dos concluintes de computação usando técnica de agrupamento. In *Anais do XXXIII Simpósio Brasileiro de Informática na Educação*, pages 834–845, Porto Alegre, RS, Brasil. SBC.
- [Webb et al. 2016] Webb, G., Hyde, R., Cao, H., Nguyen, H.-L., and Petitjean, F. (2016). Characterizing concept drift. *Data Mining and Knowledge Discovery*, 30.