



# **DESENVOLVIMENTO DE UM LEARNING ANALYTICS DASHBOARD (LAD) A PARTIR DE MODELOS DE MINERAÇÃO DE DADOS EDUCACIONAIS**

**Trabalho de Conclusão de Curso**

**Engenharia da Computação**

**Gabriel Lenon Barros da Silva**

**Orientador: Prof. Dr. Alexandre Magno Andrade Maciel**



**Universidade de Pernambuco  
Escola Politécnica de Pernambuco  
Graduação em Engenharia de Computação**

**GABRIEL LENON BARROS DA SILVA**

**DESENVOLVIMENTO DE UM  
LEARNING ANALYTICS DASHBOARD  
A PARTIR DE MODELOS DE  
MINERAÇÃO DE DADOS  
EDUCACIONAIS**

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia de Computação pela Escola Politécnica de Pernambuco – Universidade de Pernambuco.

**Recife, novembro de 2020.**

Gabriel Lenon Barros da Silva

Desenvolvimento de um Learning Analytics Dashboard (LAD) a partir de Modelos de Mineração de Dados Educacionais / Gabriel Lenon Barros da Silva – Recife – PE, Brasil, novembro 2020-

54 p.

Orientador: Alexandre Magno Andrade Maciel

Trabalho de Conclusão de Curso – Engenharia da Computação

Escola Politécnica de Pernambuco

Universidade de Pernambuco, novembro de 2020

1. Engenharia de Software 2. Learning Analytics Dashboard 3. Ferramenta para visualização de dados educacionais I. Prof. Alexandre Magno Andrade Maciel. II. Universidade de Pernambuco. III. Escola politécnica, IV. Título

## MONOGRAFIA DE FINAL DE CURSO

### Avaliação Final (para o presidente da banca)\*

No dia 12/11/2020, às 15h00min, reuniu-se para deliberar sobre a defesa da monografia de conclusão de curso do(a) discente **GABRIEL LENON BARROS DA SILVA**, orientado(a) pelo(a) professor(a) **ALEXANDRE MAGNO ANDRADE MACIEL**, sob título Desenvolvimento de um Learning Analytics Dashboard (LAD) a partir de Modelos de Mineração de Dados Educacionais, a banca composta pelos professores:

**JORGE CAVALCANTI BARBOSA FONSECA**

**ALEXANDRE MAGNO ANDRADE MACIEL (ORIENTADOR) (PRESIDENTE)**

Após a apresentação da monografia e discussão entre os membros da Banca, a mesma foi considerada:

Aprovada

Aprovada com Restrições\*

Reprovada

e foi-lhe atribuída nota: 9. ( nove e )

\*(Obrigatório o preenchimento do campo abaixo com comentários para o autor)

O(A) discente terá \_\_\_\_\_ dias para entrega da versão final da monografia a contar da data deste documento.



AVALIADOR 1: Prof (a) **JORGE CAVALCANTI BARBOSA FONSECA**



AVALIADOR 2: Prof (a) **ALEXANDRE MAGNO ANDRADE MACIEL**

AVALIADOR 3: Prof (a)

\* Este documento deverá ser encadernado juntamente com a monografia em versão final.

*Dedico este trabalho aos meus pais Ana e Claudio, por sempre me darem suporte e o conforto necessário em todos os momentos. Também dedico a Rebeka, que sempre esteve ao meu lado durante todo o desenvolvimento do trabalho. A minha vó Dora e a todos que sempre me apoiaram e encorajaram.*

# AGRADECIMENTOS

Agradeço primordialmente a Deus, Aquele que sempre me concedeu a capacidade de obter e compartilhar conhecimento. Agradeço ao meu orientador, Alexandre Maciel, por todos os conselhos, motivação e ajuda nesse período. A Raniel e Janaína, que foram fundamentais para a realização deste trabalho. A todos os membros do grupo de pesquisa GPDCA, que contribuíram com percepções essenciais para o direcionamento da pesquisa.

## Autorização de publicação de PFC

Eu, **Gabriel Lenon Barros da Silva** autor(a) do projeto de final de curso intitulado: **Desenvolvimento de um Learning Analytics Dashboard (LAD) a partir de Modelos de Mineração de Dados Educacionais**; autorizo a publicação de seu conteúdo na internet nos portais da Escola Politécnica de Pernambuco e Universidade de Pernambuco.

O conteúdo do projeto de final de curso é de responsabilidade do autor.

*Gabriel Lenon Barros da Silva*

\_\_\_\_\_  
**Gabriel Lenon Barros da Silva**



\_\_\_\_\_  
Orientador(a): **Alexandre Magno Andrade Maciel**

\_\_\_\_\_  
Coorientador(a):



\_\_\_\_\_  
Prof. de TCC: **Daniel Augusto Ribeiro Chaves**

\_\_\_\_\_  
Data: 12/11/2020

# RESUMO

Um *Learning Analytics Dashboard* (LAD) possibilita aos professores monitorarem a situação de seus alunos e receber informações úteis e de fácil entendimento sobre os mesmos. Desse modo, este trabalho propõe uma arquitetura de LAD para análise de dados a partir dos resultados de mineração de dados educacionais, através da predição do desempenho de alunos. Assim, o objetivo geral deste trabalho é a criação de um mecanismo de visualização que pretende permitir aos professores identificar os alunos dos quais o desempenho predito é não satisfatório, fornecendo informações relativas à trajetória dos alunos no curso e assim possibilitando ao professor a intervenção pedagógica e motivacional. Durante o processo de desenvolvimento da arquitetura, três estágios foram considerados: consciência, onde o foco é a visualização dos dados de diferentes formas; reflexão, cujo objetivo é avaliar a relevância dos dados presentes nas visualizações, por meio de questões feitas pelos próprios professores; *sensemaking*, que diz respeito à análise e reflexão das respostas dadas às perguntas feitas, gerando novos *insights*. Com isso, foi desenvolvida uma ferramenta cujo objetivo é apresentar predições de desempenho de alunos de uma forma amigável para os professores, de modo a contribuir com o processo de ensino-aprendizagem e potencializar a interação entre alunos, professores e recursos em ambientes virtuais de aprendizagem.

**Palavras-chave:** Análise de dados, Mineração de dados educacionais, Desempenho, *Learning Analytics Dashboard*, Ambientes virtuais de aprendizagem.



# ABSTRACT

A learning analytics dashboard (LAD) allows teachers to monitor their students' situation and receive useful and easy-to-understand information about them. This work proposes a LAD architecture for data analysis from the results of educational data mining, through the prediction of student performance. Thus, the general objective of this work is to create a visualization mechanism that aims to allow teachers to identify students whose predicted performance is unsatisfactory, providing information regarding the trajectory of students in the course and hence enabling the teacher to a pedagogical and motivational intervention. During the architecture development process, four stages were considered: awareness, where the focus is on visualizing data in different ways; reflection, whose objective is to assess the relevance of the data present in the visualizations, through questions asked to the teachers; sensemaking, which concerns the analysis and reflection of the answers given to the questions asked, creating new insights; impact, the final goal, which can produce changes in the course's conduction by teachers, if they deem it necessary. Thereby, a tool was developed whose objective is to present student performance predictions in a friendly way for teachers, in order to contribute to the teaching-learning process and empower the interaction between students, teachers and resources in learning management systems.

**Keywords:** Learning Analytics. Educational Data Mining. Performance. Learning Analytics Dashboard. Learning Management System.

# ÍNDICE DE FIGURAS

<b>Figura 1.</b> Elementos do ensino a distância	10
<b>Figura 2.</b> Modelo de Processo de Learning Analytics	12
<b>Figura 3.</b> Processo de comunicação do LAD	23
<b>Figura 4.</b> Tela de Filtros - LAD do FMDEV	26
<b>Figura 5.</b> Visualização inicial do LAD - Filtros	29
<b>Figura 6.</b> Tela de Visão Geral do LAD (1)	30
<b>Figura 7.</b> Tela de Visão Geral do LAD (2)	31
<b>Figura 8.</b> Variação do gráfico “Situação atual dos alunos” em uma turma onde não há avaliação de webquests	31
<b>Figura 9.</b> Tela de Alunos do LAD - aluno cujo desempenho predito é satisfatório	32
<b>Figura 10.</b> Tela de Alunos do LAD - aluno cujo desempenho predito é insatisfatório	33
<b>Figura A.1</b> Tela de Fonte de Dados do FMDEV	42
<b>Figura A.2</b> Tela de Pré-processamento dos dados do FMDEV	42
<b>Figura A.3</b> Tela de análise do Pré-processamento dos dados do FMDEV	43
<b>Figura A.4</b> Tela de Modelos Salvos do FMDEV	43

# ÍNDICE DE TABELAS

<b>Tabela 1.</b> Learning e Academic Analytics	6
<b>Tabela 2.</b> Distribuição de alunos por disciplina	14
<b>Tabela 3.</b> Indicadores dos alunos	15-17
<b>Tabela 4.</b> Modelos representados por construtos	18-20
<b>Tabela 5.</b> Visualizações implementadas no LAD	27-28
<b>Tabela 6.</b> Questionário de validação do LAD respondido por professores	35-36

# TABELA DE SÍMBOLOS E SIGLAS

AVA	Ambiente Virtual de Aprendizagem
EAD	Ensino a Distância
MDE	Mineração de Dados Educacionais
FMDEV	Framework de Mineração de Dados Educacionais
LA	<i>Learning Analytics</i>
LAD	<i>Learning Analytics Dashboard</i>

# SUMÁRIO

<b>INTRODUÇÃO</b>	<b>1</b>
<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>3</b>
Ambientes Virtuais de Aprendizagem	3
Mineração de Dados Educacionais	4
<i>Learning Analytics e Visualização de Dados</i>	5
<i>Learning Analytics</i>	5
<i>Learning Analytics Dashboard</i>	7
Princípios do Design de Dashboard	8
Desempenho do Aluno no EAD	9
<b>METODOLOGIA</b>	<b>12</b>
Dados e Ambiente Utilizados	13
Modelos Treinados	17
Desenvolvimento das Visualizações	21
Validação da Ferramenta	21
<b>SOLUÇÃO PROPOSTA</b>	<b>23</b>
Constituição da ferramenta	23
Tecnologias Utilizadas	23
Ambiente de desenvolvimento	25
Design do LAD	27
<b>VALIDAÇÃO E RESULTADOS</b>	<b>34</b>
<b>CONCLUSÃO E TRABALHOS FUTUROS</b>	<b>38</b>
<b>ANEXOS</b>	<b>42</b>
<b>ANEXO A - TELAS DO FMDEV</b>	<b>42</b>

# CAPÍTULO 1

## INTRODUÇÃO

Por conta dos avanços tecnológicos, o ensino a distância tem se tornado uma modalidade comum e cada vez mais útil nos ambientes universitários. Ao lado disso, acompanha o crescimento da quantidade de dados gerados por alunos em Ambientes Virtuais de Aprendizagem (AVA), plataformas que auxiliam no acompanhamento e desenvolvimento dos alunos e que servem de intermediação no ensino-aprendizado destes.

Muito utilizado para conduzir o ensino a distância, o AVA também tem sido útil no apoio da educação presencial, visto a sua capacidade de, por meio de ferramentas, prover interação entre alunos, professores e recursos educacionais.

Com essa frequente interação entre os usuários no AVA, vê-se a necessidade da aplicação de *Learning Analytics* (LA), que, por meio da medição, coleta, análise e divulgação de dados sobre os alunos em diferentes contextos, tem o propósito de entender e aperfeiçoar a aprendizagem e os ambientes nos quais ela ocorre (Long e Siemens, 2011).

A partir do uso de LA na compreensão de dados educacionais para o sucesso da aprendizagem (Rosales, 2014), surge a problemática de se encontrar dados que devam ser considerados na análise, a tradução desses dados em representações visuais que permitam oferecer feedback para apoio da aprendizagem, o aspecto da privacidade necessária aos traços registrados pelos estudantes e possibilitar melhorar a aprendizagem e minimizar a reprovação dos alunos (Verbert et al., 2013).

Através dessa perspectiva, viu-se a oportunidade de criar uma ferramenta para análise de dados a partir dos resultados de mineração de dados educacionais, definida como um *Learning Analytics Dashboard*, uma ferramenta que pode dar suporte ao ensino e ao aprendizado que funciona através do rastreamento de arquivos de registro dos alunos e do fluxo de dados em tempo real, extraíndo grandes quantidades de dados para encontrar significado e gerando visualização

para os resultados, para que possam ser compreendidos facilmente à primeira vista (Park e Jo, 2015).

Do ponto de vista arquitetural, será desenvolvida uma ferramenta que deve representar um novo módulo do FMDEV, um ambiente de Auto ML (aprendizado de máquina automatizado) cujo objetivo é permitir que usuários, quer sejam técnicos ou não, possam construir, validar e disponibilizar baselines de AM (Aprendizado de máquina) com maior produtividade e com menor conhecimento em ciência de dados (Silva, 2020).

# CAPÍTULO 2

## 2. FUNDAMENTAÇÃO TEÓRICA

Nessa seção são apresentados os conceitos e fundamentos necessários para a compreensão, o acompanhamento, e o próprio desenvolvimento do trabalho. O objetivo aqui é discorrer sobre os trabalhos relacionados e buscar facilitar o entendimento do projeto através da definição dos seguintes conceitos: Ambientes Virtuais de Aprendizagem, Mineração de Dados Educacionais, *Learning Analytics* e Visualização de dados, Desempenho do Aluno no Ensino a Distância.

### 2.1. Ambientes Virtuais de Aprendizagem

Ambientes Virtuais de Aprendizagem (AVA) são o resultado da evolução do Ensino a Distância (EAD), já que, desde o surgimento do EAD em 1728, com a disponibilização de um curso de taquigrafia que era enviado semanalmente pelos correios pelo professor Caleb Phillips (Litto e Formiga, 2012), e a partir da integração de computadores no processo de ensino-aprendizagem, até os dias de hoje, viu-se um avanço através da utilização de novos recursos no EAD que possibilitam o acesso a informações sistematizadas e uma interação entre os alunos e os professores (Faria e Salvadori, 2010).

Muitos autores definem o que são AVA's, sempre relacionados a sistemas de software, mas é imprescindível mencionar que são concebidos para apoiar o processo de ensino e ajudar os alunos a aprenderem por meio do gerenciamento da aprendizagem (Kats, 2013), sendo possível, através deles, a realização de coletas de dados sobre as interações que os estudantes têm entre si, com os professores e com recursos educacionais, sendo estes dados indicadores valiosos sobre a eficácia de uma experiência de aprendizagem (Romero-Zaldivar et al., 2012).

A partir do armazenamento e gerenciamento dos dados mencionados, tais como registros de acessos, visualização dos materiais das disciplinas, participação em fóruns, resolução de questionários e envio de atividades, notas dos alunos e o horário em que os eles acessam o sistema, dentre outros, é possível se fazer



análises sobre os elementos que podem influenciar o processo de ensino-aprendizagem e, conseqüentemente, o desempenho dos alunos na modalidade do ensino a distância.

## 2.2. Mineração de Dados Educacionais

A Mineração de Dados (MD) é definida como um processo de descoberta de padrões úteis ou conhecimento a partir de fontes de dados (Liu, 2016), podendo os dados se apresentarem estruturados, semi-estruturados ou não estruturados.

Tendo isso em vista, diferentes métodos podem ser aplicados para minerar os dados, sendo os métodos separados em categorias, sendo as que seguem tidas como as principais: aprendizagem supervisionada, na qual é feita a classificação e a previsão do dado; aprendizagem não supervisionada, em que ocorre o agrupamento de dados por características semelhantes; regras de associação (Liu, 2006).

A aplicação de técnicas de mineração de dados objetiva dar suporte e atendimento às diferentes necessidades de usuários, viabilizando a construção de ferramentas analíticas a partir de modelos de análise preditivos e prescritivos (Rosales, 2014), podendo atuar em diferentes âmbitos, como o educacional, esfera de atuação do presente trabalho.

Nesse contexto, é introduzido o conceito de Mineração de Dados Educacionais (MDE), que faz o uso de técnicas computacionais e estatísticas para visualização de dados educacionais (Romero e Ventura, 2007).

Ademais, conforme é dito por Rosales (2014), os conhecimentos já produzidos na área de MDE, especialmente a adoção de técnicas de mineração de dados, são utilizados a fim de atingir melhores resultados de análises. Desse modo, vê-se como promissora a junção da aplicabilidade da mineração com a visualização de dados educacionais, podendo gerar novos *insights* no decorrer dos cursos em ambientes virtuais de aprendizagem.

## 2.3. Learning Analytics e Visualização de Dados

Essa seção abordará sobre os seguintes conceitos: *Learning Analytics*, *Learning Analytics Dashboard* e Princípios do Design de *dashboard*.

### 2.3.1. Learning Analytics

Com o crescente número de dados que têm sido produzidos na área da educação, faz-se necessário uma forma de gerenciar e analisar esses dados, para que os mesmos possam ser utilizados em prol do constante desenvolvimento do ensino. Quando se fala de ensino a distância e ambientes virtuais de aprendizagem, a aplicabilidade desses dados pode ser potencializada, visto que, utilizando-se *Learning Analytics* (LA), é possível prover ferramentas educacionais poderosas para apoiar os usuários de educação online, oferecendo melhores condições de sucesso para aprendizes, além de melhorar a efetividade das práticas pedagógicas e as tomadas de decisão de instrutores, no âmbito institucional (Rosales, 2014).

LA é uma área emergente que explora a medição, coleta, análise e relatório de dados que estão associados à aprendizagem dos alunos e seu contexto, com o propósito de entender e aperfeiçoar a aprendizagem e os ambientes nos quais ela ocorre (Brown, 2011). Siemens e Long (2011) acrescentam que, através desse processo realizado sobre os dados, pode-se observar e entender seus comportamentos para possibilitar intervenção apropriada. Ainda para Gilfus (2011), LA é o uso de dados e modelos para prever o progresso e o desempenho do estudante, e a possibilidade de agir sobre as informações.

É importante evidenciar que o objeto de pesquisa do presente trabalho, no ambiente de LA, tem foco mais exclusivo no processo de aprendizagem, que envolve o relacionamento entre aluno, conteúdo, instituição e professor, enquanto *Academic analytics*, apesar de ser similar, reflete o papel da análise de dados em um nível institucional (Siemens e Long, 2011). Tais diferenças são melhor detalhadas na Tabela 1.

**Tabela 1.** Learning e Academic Analytics [adaptado de Siemens e Long, 2011]

Tipo de Analytics	Nível ou objeto de análise	Quem se beneficia?
Learning Analytics	<b>A nível de curso:</b> redes sociais, desenvolvimento conceitual, análise de discurso, “currículo inteligente”	Alunos, corpo docente
	<b>Departamental:</b> modelagem preditiva, padrões de sucesso/falha	Alunos, corpo docente
Academic Analytics	<b>Institucional:</b> perfis de alunos, desempenho de acadêmicos, fluxo de conhecimento	Administradores, financiadores, marketing
	<b>Regional</b> (estatal/provincial): comparações entre sistemas	Financiadores, administradores
	<b>Nacional e Internacional</b>	Governos nacionais, autoridades da educação

Segundo Park e Jo (2015), dentre as diversas abordagens para tais análises de aprendizagem, o uso de dados de registro em AVA's é uma das orientações de pesquisa mais populares devido a sua onipresença em muitas instituições educacionais.

Nesse contexto, com o uso de LA para análise e interpretação automática dos dados, é possível fornecer ao professor informações que possam ser entendidas à primeira vista, auxiliando-o no entendimento das necessidades de aprendizagem dos estudantes. Desse modo, a aplicação de LA no contexto educacional tem implicações tanto para o desempenho individual dos estudantes como para os processos de ensino, aprendizagem e avaliação.

### 2.3.2. *Learning Analytics Dashboard*

*Dashboards* normalmente tem por objetivo capturar e apresentar, de forma visual, o rastreamento das atividades de aprendizagem com a finalidade de promover a consciência, reflexão e permitir que os estudantes possam definir metas e acompanhar seu progresso na obtenção desses objetivos (Verbert, et al., 2014).

De acordo com Li, Dey, e Forlizzi (2010) *Learning Dashboards* podem ser considerados como uma classe específica de aplicações de informática pessoal, que usualmente ajudam usuários na coleta de informações pessoais a respeito de diversos aspectos de suas vidas, comportamentos, hábitos, ideias e interesses. Assim, *dashboards* desenvolvidos para diversos propósitos dão suporte a professores, alunos ou ambos em uma abordagem sinérgica que combina princípios e tecnologias de *design*, incluindo extração de dados, mineração de dados e visualização (Park e Jo, 2015).

Park e Jo (2015) caracterizam *Learning Analytics Dashboard* (LAD) como uma aplicação para mostrar aos estudantes padrões de comportamento em um ambiente virtual de aprendizagem. E é a partir desses padrões que é possível se obter resultados úteis no ambiente educacional, visto que o objetivo de um *dashboard* é indicar as informações mais importantes que são necessárias para se alcançar um ou mais objetivos, através de uma única tela de computador, para que tais informações possam ser monitoradas à primeira vista (Few, 2013).

Finalmente, os LAD's propriamente ditos, que representam o modelo prescritivo de LA, retêm papéis e valores únicos (Park e Jo, 2019). Desse modo, Park e Jo (2019) acrescentam que um *dashboard*, quando eficaz, influencia o comportamento auto regulatório dos alunos e, em última análise, seus comportamentos de aprendizagem e resultados.

A finalidade da pesquisa é desenvolver um LAD para uso de professores, com o objetivo de informar sobre o estado da aprendizagem do estudante, um dos tipos de *dashboard* de aprendizagem destacados por Park e Jo (2015), que também enfatizam para outros dois tipos de *dashboard* de aprendizagem: *dashboard* para estudantes, que apresentam padrões aos estudantes, ajudando-os a modificar suas estratégias, motivando-os a aprendizagem e a autorreflexão; *dashboards* para apoio

a professores e estudantes, que reúnem características comuns aos dois tipos de *dashboards* citados anteriormente.

### 2.3.3. Princípios do Design de Dashboard

No contexto educacional, o objetivo dos *dashboards* é o facilitar o entendimento de registros de dados de alunos, tanto a respeito de seus comportamentos no contexto ao qual eles estão inseridos, como da interação com os recursos educacionais presentes nos ambientes de aprendizagem. Tais *dashboards*, quando evidenciam o sentido e a significância dos dados, podem, através de visualização simples e objetiva, permitir aos professores analisar informações relativas aos alunos e contribuir na tomada de decisão durante o processo de ensino-aprendizagem.

Para alcançar tais objetivos, Few (2013) ressalta a importância do desenvolvimento de *dashboards* bem projetados, que levam a uma comunicação eficaz, bem como à tomada de decisão correta.

No que diz respeito à percepção visual, Few (2013) faz três considerações notáveis da literatura:

1. Os seres humanos têm memória de trabalho limitada; apenas três ou quatro blocos de informações visuais podem ser armazenados por vez. Portanto, para projetar um *dashboard*, faz-se necessário exibições bem projetadas como gráficos, pois são melhores para percepção eficiente e retenção de memória, ao contrário de números individuais.
2. Para uma percepção rápida, atributos pré-ativos como cor, forma, posição espacial e movimento devem ser utilizados de maneira adequada.
3. Seguindo os princípios da Gestalt, elementos como proximidade, similaridade, enclausuramento, fechamento, continuidade e conexão devem ser considerados na concepção de design de um *dashboard*.

A partir das considerações feitas por Few, Park e Jo (2015) resumem as implicações destes princípios no design de *dashboards*:

1. A informação mais importante deve se destacar das demais no *dashboard*, que normalmente tem um espaço limitado;

2. As informações no *dashboard* devem apoiar a consciência situada e ajudar a percepção rápida usando diversas tecnologias de visualização;
3. As informações devem ser dispostas de forma que façam sentido, e os elementos das informações devem contribuir para os objetivos imediato e final do visualizador, para a tomada de decisão.

## 2.4. Desempenho do Aluno no EAD

Normalmente, os LAD's têm, por objetivo inicial, contribuir com o processo de ensino-aprendizagem, e por objetivo final, a otimização do desempenho dos alunos.

Siemens e Long (2011) destacam que para educadores, a disponibilidade de informações em tempo real sobre o desempenho dos alunos, incluindo os que estão em risco, pode ser uma ajuda significativa no planejamento das atividades de ensino. Para os alunos, receber informações sobre seu desempenho em relação aos colegas ou sobre sua evolução em relação a seus objetivos pessoais pode ser motivador e encorajador.

Nesse cenário, surge a importância de buscar entender quais os fatores que afetam o desempenho do aluno no ensino a distância, tanto positivamente como negativamente.

Shih et al. (2003) afirma que os elementos da educação a distância, que podem ser divididos em três categorias, conforme ilustrado na Figura 1, são componentes essenciais que afetam o desenvolvimento de programas de ensino a distância.



**Figura 1.** Elementos do ensino a distância [adaptado de Shih et al., 2003]

Por se utilizar essencialmente de recursos tecnológicos, o ensino a distância traz os seus obstáculos e dificuldades que podem afetar diretamente o desempenho de alguns alunos que eventualmente não tenham tido uma experiência anterior com esses recursos, que são necessários no dia a dia da modalidade EAD.

Com base nisso, Rabello (2007) organiza as dificuldades do aluno no ensino a distância em duas categorias: dificuldades pessoais, que diz respeito à metodologia de educação a distância, e dificuldades com o desenho instrucional, que diz respeito às dificuldades relacionadas ao modelo instrucional do curso. Já Ramos (2013) agrupa-as em três categorias: dificuldades cognitivas, pedagógicas e pessoais.

De acordo com Ramos et.al (2014), os fatores que podem influenciar o desempenho dos alunos nessa modalidade envolvem novas variáveis que não são observadas no ensino presencial, como conhecimento tecnológico prévio, tempo de acesso no ambiente virtual, apoio de tutoriais e material didático do curso.

Este trabalho se utiliza da abordagem de uma outra importante vertente na pesquisa do desempenho de estudantes na EAD, que trata do estabelecimento de análises e modelos de previsão de desempenho dos aluno, utilizando de forma cada

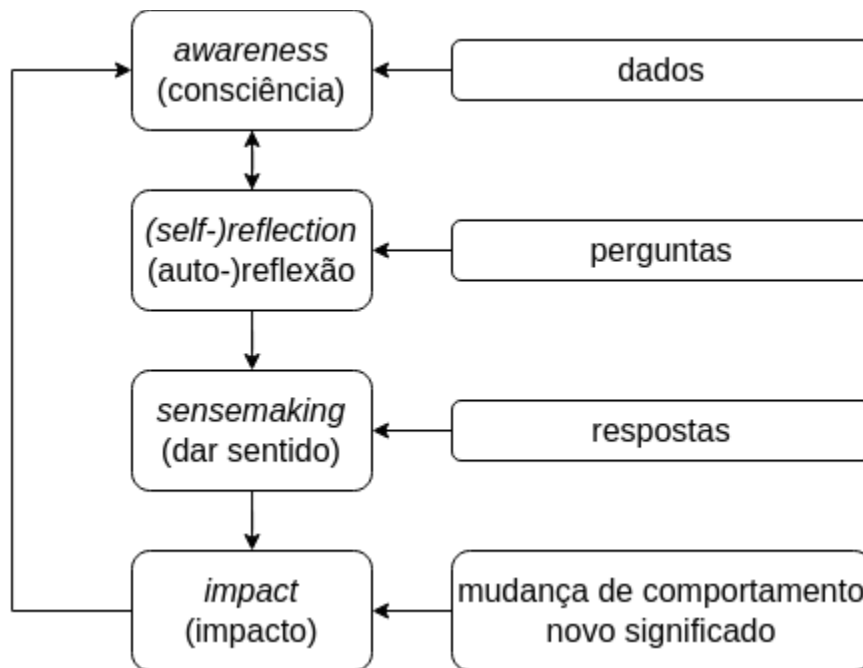
vez mais consistente a Mineração de Dados Educacionais (Ramos et.al, 2014). Ramos acrescenta que, a partir das diversas interações e registros de acesso dos alunos nos ambientes virtuais, é possível descrever um modelo que permite medir ou prever desempenho de alunos em um curso online.



# CAPÍTULO 3

## 3. METODOLOGIA

A metodologia utilizada neste trabalho foi fundamentada a partir do modelo de processo proposto por Verbert et al. (2013), onde se afirma que aplicações de LA podem ser distinguidas em quatro fases, como pode ser visto na Figura 2, sendo elas: *awareness* (consciência), *reflection* (reflexão), *sensemaking* (dar sentido) e *impact* (impacto).



**Figura 2.** Modelo de Processo de Learning Analytics [adaptado de Verbert et al., 2013]

A partir desse modelo de processo, foi feita uma adaptação para que três dos quatro estágios fossem considerados, sendo eles: *awareness*, *reflection* e *sensemaking*. O quarto estágio, referente ao impacto, ficou destacado para que seja implementado em um trabalho futuro, visto que, nesse estágio, é preciso fazer uma análise de possíveis novos *insights* gerados pelos usuário finais (professores), podendo provocar ação e mudança de comportamento (Verbert et al., 2013). Para

isso, seria preciso o acompanhamento da abordagem de ensino de um professor para com sua turma.

Com o modelo de processo adaptado, os seguintes passos foram percorridos durante o desenvolvimento do LAD: entendimento dos dados utilizados e do contexto/ambiente para estudo de caso; treinamento de modelos de *machine learning* para predição do desempenho de alunos, baseados em construtos propostos pelo modelo teórico de Pintrich (1999) e nas variáveis da base de dados do NEAD - UPE (núcleo de ensino a distância da UPE), no AVA Moodle, relativas ao processo de aprendizagem autorregulatória dos alunos (Rodrigues et al., 2019); desenvolvimento de visualizações para representar os dados preditivos e descritivos (*awareness*), a partir do qual os usuários refletem a respeito dos dados e refletem sobre si mesmos para avaliar seu comportamento anterior (*self-reflection*), e posteriormente os usuários respondem às suas próprias perguntas e criam novos *insights* (*sensemaking*). Nessa fase, houve a elaboração de perguntas para serem respondidas por meio de questionário aplicado a professores do ensino superior, depois dos mesmos utilizarem o LAD. Com isso, foi possível obter as respostas do questionário para avaliação da utilidade, facilidade e intenção de uso percebidas, obtendo-se uma validação da aceitação da ferramenta.

### 3.1. Dados e Ambiente Utilizados

Uma das coisas mais importantes, ao desenvolver uma ferramenta de LA e visualização, é o entendimento dos dados que serão utilizados, o seu contexto e o ambiente onde esses dados são tratados, além de onde se desenvolve a aplicação.

No que se refere a coleta de dados, foi disponibilizado pelo NEAD - UPE um arquivo no formato .csv referente ao backup da sua base de dados no Moodle, contendo dados de logs de alunos e os seus indicadores no AVA mencionado. Esses dados são relativos às disciplinas de Administração, Biologia, Letras e Pedagogia, cursadas no período de 2009 a 2016.

Para utilização dos dados no mesmo ambiente do LAD, foi feita a importação do arquivo .csv em uma tabela que foi criada na base de dados do FMDEV, tabela essa que foi chamada de “moodle”.

A tabela “moodle” possui 30.217 registros e uma de suas colunas diz respeito a disciplina cursada, a partir da qual é possível identificar a distribuição de alunos por disciplina, como pode ser visto na Tabela 2.

**Tabela 2.** Distribuição de alunos por disciplina

<b>Disciplina</b>	<b>Quantidade de alunos</b>
Administração	2.892
Biologia	6.526
Letras	6.297
Pedagogia	14.502
Total	30.217

Os dados utilizados no desenvolvimento do LAD, delineando um estudo de caso, são uma parte importante na definição do contexto, o qual atua como um conjunto de informações que influenciam o comportamento do sistema (Bazire e Brézillon, 2005). Com isso, a análise do contexto é importante para o presente trabalho, posto que, de acordo com Yin (2015), o estudo de caso investiga um fenômeno contemporâneo em seu contexto no mundo real.

Outra parte importante na definição do contexto, no qual o LAD é desenvolvido, são os indicadores que representam o comportamento, as ações e as interações dos alunos no AVA Moodle. Os indicadores dos alunos são representados por colunas na tabela “moodle”, e o significado de cada indicador é descrito na tabela 3.

Tabela 3. Indicadores dos alunos [adaptado de Rodrigues et al., 2019]

<b>Indicador</b>	<b>Descrição</b>
var01	Número de diferentes lugares de onde o aluno acessou o AVA.
var02	Número de mensagens enviadas do aluno ao(s) professor(es) usando o AVA.
var03	Número de mensagens enviadas do aluno ao(s) tutor(es) usando o AVA.
var04	Número total de mensagens enviadas pelo aluno no AVA.
var05	Número total de mensagens recebidas pelo aluno no AVA.
var06	Número de threads criadas pelo aluno em fóruns Q&A.
var07	Número de threads no fórum Q&A.
var08	Número de threads no fórum criadas pelo aluno que foram respondidas por outros alunos.
var09	Número de threads no fórum criadas pelo aluno que foram respondidas pelo tutor ou professor.
var10	Número de diferentes colegas para quem o aluno enviou mensagens no AVA.
var12	Número de vezes que a seção "Conteúdos" (que lista os arquivos que descrevem o programa do curso) foi visualizada.
var13	Hora do dia em que o aluno mais frequentemente trabalha em suas atribuições/exercícios.
var14	Período do dia (manhã, tarde, anoitecer, noite) na qual o aluno mais frequentemente trabalhou em suas atribuições/exercícios.

var16	Número de atribuições/exercícios entregues pelo aluno após o prazo.
var17	Tempo médio entre o momento em que uma atividade foi atribuída e em que o aluno a completou.
var18	Número de vezes que o aluno acessa o fórum (pageviews).
var20	Número de respostas em uma thread no fórum (denota a ação de reconsiderar a opinião sobre o assunto).
var21	Número de vezes que o aluno acessa o relatório de notas.
var22	Número de vezes que o aluno visualizou as atividades.
var23	Número de visualizações nas notas de atividades.
var24	Média semanal do número de vezes em que o aluno acessou o AVA.
var25	Tempo médio entre o momento em que um tópico é criado no fórum e o momento em que o aluno posta sua primeira resposta nele.
var28	Número de timeouts no AVA.
var31	Número de vezes que o aluno acessou o AVA.
var31b	Número de distintos dias onde o aluno acessou o curso no AVA.
var31c	Número de distintos dias onde o aluno acessou o AVA.
var32a	Número de vezes em que o aluno acessou o AVA por período do dia (manhãs).
var32b	Número de vezes em que o aluno acessou o AVA por período do dia (tardes).

var32c	Número de vezes em que o aluno acessou o AVA por período do dia (fins de tardes).
var32d	Número de vezes em que o aluno acessou o AVA por período do dia (noites).
var33	Número de atividades/atribuições entregues por um aluno dentro do prazo.
var34	Número total de mensagens postadas pelo aluno nos fóruns.
var35	Número de respostas de um professor para perguntas do aluno em fóruns.

Além dos indicadores descritos na Tabela 2, também estão disponíveis outras colunas na tabela “moodle”, das quais as que seguem foram consideradas importantes nas análises criadas no LAD: curso, semestre, período, nome da disciplina, nome do aluno, primeira prova, segunda prova, média das provas, fórum 01, fórum 02, fórum 03, fórum 04, média dos fóruns, webquest (atividade de pesquisa) 01, webquest 02, média dos webquests e desempenho binário, sendo este último o indicador primário na análise de desempenho dos alunos, visto que representa a reprovação ou a aprovação do aluno na disciplina, podendo possuir os valores 0 e 1, respectivamente.

## 3.2. Modelos Treinados

Um dos objetivos do trabalho é a criação de uma aplicação que possa permitir aos professores identificar os alunos dos quais o desempenho predito é não satisfatório, fornecendo informações relativas à trajetória dos alunos no curso, representadas pelos indicadores que estão dispostos nas colunas da tabela “moodle”.

Para que fosse possível fazer a predição do desempenho dos alunos, modelos de *machine learning*, cujo indicador alvo é o desempenho binário, foram treinados no ambiente do FMDEV.

No treinamento dos modelos, foram utilizados os indicadores que estão descritos na Tabela 2, uma vez que os mesmos são dados de logs que podem representar construtos relacionados ao processo de autorregulação do aprendizado dos alunos (Rodrigues et al., 2019). Desse modo, conforme proposto por Rodrigues et al., (2019), os indicadores foram agrupados de modo a compor seis construtos de autorregulação da aprendizagem, os quais foram utilizados como base da geração dos modelos de *machine learning*.

Na Tabela 4 são mostrados os construtos que representam os modelos de *machine learning* que foram treinados no FMDEV e ficaram disponíveis para uso no LAD, tal como os indicadores que fazem parte de cada construto, sendo utilizados tanto no pré-processamento como treinamento dos modelos.

**Tabela 4.** Modelos representados por construtos [adaptado de Rodrigues et al., 2019]

<b>Construto</b>	<b>Indicador</b>	<b>Descrição</b>
Estruturação do Ambiente	var01	Número de diferentes lugares (Endereços IP) de onde o aluno acessou o AVA.
Em busca de ajuda	var02	Número de mensagens enviadas do aluno ao(s) professor(es) usando o AVA.
	var03	Número de mensagens enviadas do aluno ao(s) tutor(es) usando o AVA.
	var04	Número total de mensagens enviadas pelo aluno no AVA.
	var05	Número total de mensagens recebidas pelo aluno no AVA.

	var06	Número de threads criadas pelo aluno em fóruns Q&A.
	var07	Número de threads no fórum Q&A.
	var08	Número de threads no fórum criadas pelo aluno que foram respondidas por outros alunos.
	var09	Número de threads no fórum criadas pelo aluno que foram respondidas pelo tutor ou professor.
	var10	Número de diferentes colegas para quem o aluno enviou mensagens no AVA.
Estratégias de conclusão de tarefas	var12	Número de vezes que a seção "Conteúdos" (que lista os arquivos que descrevem o programa do curso) foi visualizada.
	var13	Hora do dia em que o aluno mais frequentemente trabalha em suas atribuições/exercícios.
	var14	Período do dia (manhã, tarde, anoitecer, noite) na qual o aluno mais frequentemente trabalhou em suas atribuições/exercícios.
	var16	Número de atribuições/exercícios entregues pelo aluno após o prazo.
	var17	Tempo médio entre o momento em que uma atividade foi atribuída e em que o aluno a completou.
	var18	Número de vezes que o aluno acessa o fórum (pageviews).
Autoavaliação	var20	Número de respostas em uma thread no fórum (denota a ação de reconsiderar a opinião sobre o assunto).
	var21	Número de vezes que o aluno acessa o relatório de notas.



	var22	Número de vezes que o aluno visualiza as atividades.
	var23	Número de visualizações nas notas de atividades.
Gestão de tempo	var24	Média semanal do número de vezes em que o aluno acessou o AVA.
	var25	Tempo médio entre o momento em que um tópico é criado no fórum e o momento em que o aluno posta sua primeira resposta nele.
	var28	Número de timeouts no AVA.
	var31	Número de vezes que o aluno acessou o AVA.
	var31b	Número de distintos dias onde o aluno acessou o curso no AVA.
	var31c	Número de distintos dias onde o aluno acessou o AVA.
	var32a	Número de vezes em que o aluno acessou o AVA por período do dia (manhãs).
	var32b	Número de vezes em que o aluno acessou o AVA por período do dia (tardes).
	var32c	Número de vezes em que o aluno acessou o AVA por período do dia (fins de tardes).
	var32d	Número de vezes em que o aluno acessou o AVA por período do dia (noites).
Definição de metas	var33	Número de atividades/atribuições entregues por um aluno dentro do prazo.
	var34	Número total de mensagens postadas pelo aluno nos fóruns.

	var35	Número de respostas de um professor para perguntas do aluno em fóruns.
--	-------	--

### 3.3. Desenvolvimento das Visualizações

Essa é a fase inicial no processo de avaliação do sucesso de um *dashboard*, onde os dados são dispostos de forma visual, sem nenhuma implicação relevante ainda, e os usuários se tornam cientes dos dados visualizados.

Em virtude disto, foram implementadas as primeiras visualizações do LAD em forma de gráficos, tabelas e *cards*, principalmente, onde seguiu-se três princípios de design para *dashboards* indicados por Park e Jo (2019):

1. A informação mais importante deve se destacar das demais no *dashboard*, que normalmente tem um espaço limitado e deve se encaixar em uma tela única;
2. As informações no *dashboard* devem apoiar a consciência situacional e facilitar na rápida percepção, usando diversas tecnologias de visualização;
3. As informações devem ser dispostas de uma forma que faça sentido, e cada elemento deve contribuir para os objetivos imediatos e finais do usuário (visualizador), para a tomada de decisão.

### 3.4. Validação da Ferramenta

Para a validação da ferramenta, foi elaborado um questionário onde foi aplicado o modelo TAM (*Technology Acceptance Model*) e a escala Likert, cujo objetivo foi avaliar a utilidade, a facilidade e a intenção de uso do LAD.

O TAM é um modelo que desenvolve e valida novas escalas para duas variáveis específicas, utilidade percebida e facilidade de uso percebida, dois fatores que podem ser teorizados como determinantes fundamentais da aceitação do usuário (Davis, 1989). Esse modelo foi desenvolvido baseado na Teoria da Ação Racional (Fishbein e Ajzen, 1975), e sua utilização na validação do LAD se deu por

ele ter o objetivo de fornecer uma explicação sobre os fatores determinantes da aceitação de uma tecnologia específica pelos usuários finais.

A escala Likert foi desenvolvida pelo cientista Rensis Likert, onde o respondente, em cada questão do questionário proposto, diz o seu grau de concordância ou discordância sobre algo, escolhendo um valor em uma escala com cinco graus originalmente. Para o questionário aplicado neste trabalho, foram utilizados os graus: concordo plenamente, concordo, neutro, discordo e discordo plenamente.

Concluindo, a validação da ferramenta proposta foi realizada com professores do ensino superior que possuem experiência na modalidade de ensino a distância, onde primeiramente foi apresentada uma introdução inicial do LAD aos mesmos, explicando sobre os passos que deveriam ser trilhados por eles na validação: utilização do LAD, pelo tempo que achassem necessário, visualizando os dados e levantando questões a si mesmos sobre como essa ferramenta poderia contribuir para seus conceitos e/ou abordagens no seu processo de ensino; resposta ao questionário, através do qual se avaliou a utilidade, a facilidade e a intenção de uso do LAD, para que seja possível compreender a aceitação da ferramenta.

# CAPÍTULO 4

## 4. SOLUÇÃO PROPOSTA

Para o desenvolvimento da arquitetura proposta neste trabalho, foi feita uma integração interna ao FMDEV (Framework de Mineração de Dados Educacionais), de modo que um novo módulo, referente ao LAD, foi criado e integrado ao *framework*.

O objetivo da integração foi o de adicionar um módulo de visualização e análise de dados a uma aplicação de *Machine Learning*, permitindo ao professor, a partir de modelos de *Machine Learning* criados, fazer a predição do desempenho de alunos e analisar os dados relativos ao comportamento e a interação dos alunos com professores e recursos educacionais, dados estes originados de ambientes virtuais de aprendizagem.

Realizada a integração com um framework de mineração de dados, o LAD pretendeu permitir ao professor poder identificar alunos que têm alta probabilidade de reprovação, dados os indicadores relativos ao comportamento e ações do próprio aluno no AVA.

### 4.1. Constituição da ferramenta

Em relação à constituição do LAD, como dito anteriormente, foi feita a criação de um novo módulo no FMDEV, dado o propósito do mesmo de construir, validar e disponibilizar baselines de *auto machine learning* com maior produtividade e menor conhecimento em ciência de dados (Silva, 2020), e vista a necessidade de ter visualizações que possam gerar *insights* a partir da predição realizada por modelos treinados no FMDEV.

#### 4.1.1. Tecnologias Utilizadas

Como foi decidido desenvolver o LAD integrado ao FMDEV, a mesma base de *stack* de tecnologias foi utilizada. Desse modo, de forma geral, utilizou-se a linguagem Javascript no *frontend* e a biblioteca Plotly para a geração de gráficos.

Apesar de o FMDEV já apresentar uma estrutura robusta de *backend*, desenvolvido em Python, foi preciso fazer ajustes no código relativo à predição a partir de um modelo treinado, para que a API que realiza as predições recebesse os dados e os devolvesse de acordo com a necessidade do LAD.

A seguir, é descrito o que foi essencialmente utilizado para desenvolvimento do *frontend* do LAD e para os ajustes do *backend*.

- Frontend: HTML 5, CSS, Javascript, React.js, Redux e Plotly.
- Backend: Python e Flask.

Por conta de o FMDEV ter sido feito para a plataforma Web, o frontend foi desenvolvido com HTML 5, CSS e Javascript, três tecnologias que são essenciais para o desenvolvimento de qualquer aplicação desenvolvida para essa plataforma.

Além disso, seguindo a baseline de desenvolvimento do frontend do FMDEV, foi utilizado: React.js, uma biblioteca Javascript para construção de interfaces gráficas e componentes; Redux, para o gerenciamento do estado dos componentes e de toda a aplicação; Plotly, para a construção de gráficos dinâmicos, uma biblioteca que tem integração com várias linguagens de programação, inclusive Javascript.

Para os ajustes no código de predição, também seguindo a baseline de desenvolvimento do backend do FMDEV, foi utilizado o micro-framework Flask, que é desenvolvido em Python e implementa uma API REST, sendo responsável por fazer uma das partes da comunicação do frontend do LAD com o algoritmo de predição, através do protocolo HTTP.

É importante ressaltar que uma base de dados foi configurada para armazenar os dados relativos aos arquivos de registro dos alunos no AVA e fornecê-los através de consultas (queries) para o LAD. Para isso, foi criada a tabela “moodle” em um banco de dados Postgres, que também armazena as outras tabelas inerentes ao FMDEV.

Na Figura 3, pode-se ver o fluxograma do processo de comunicação do LAD, onde o usuário acessa e tem a visão inicial do frontend da aplicação através de um navegador Web, podendo, a partir daí, gerar a análise dos alunos e visualizá-la.

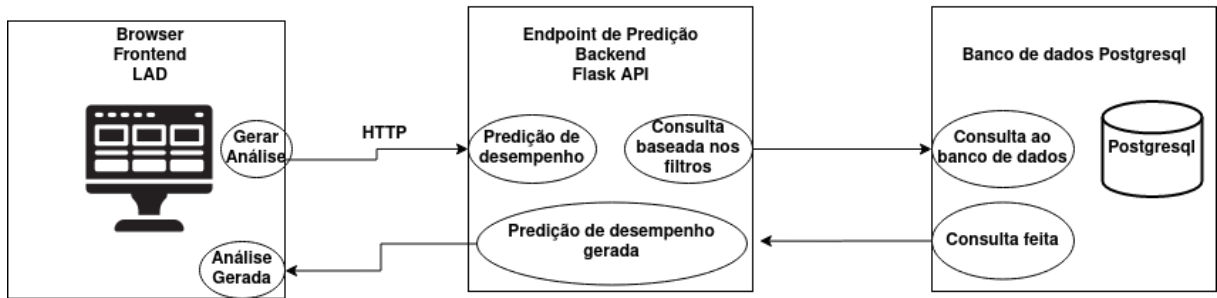
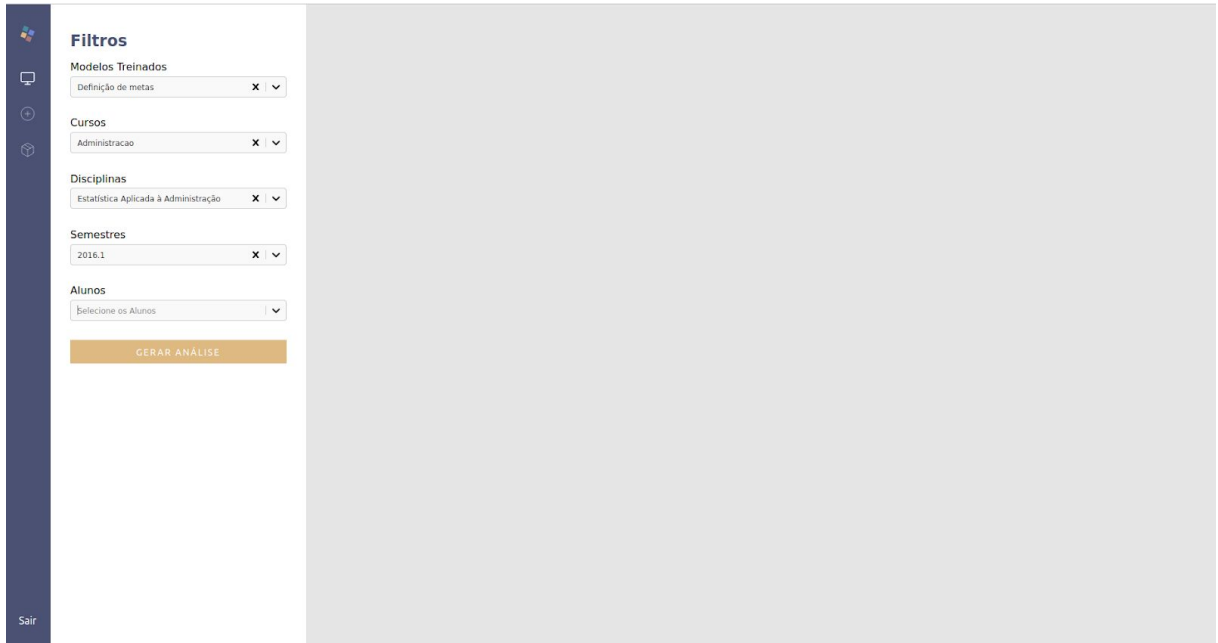


Figura 3. Processo de comunicação do LAD - [O autor, 2020]

#### 4.1.2. Ambiente de desenvolvimento

Como foi dito na subseção de tecnologias utilizadas, o ambiente onde o LAD foi desenvolvido é o mesmo do FMDEV, que, com o desenvolvimento do LAD, apresenta três telas principais:

1. **Fonte de dados:** nessa tela, o usuário pode selecionar uma fonte de dados disponibilizada pelo FMDEV ou importar a sua própria, conforme Figura A.1 do Anexo A. A partir daí, o usuário tem a opção de fazer o pré-processamento dos dados e o treinamento do modelo de *machine learning* (Figura A.2 e Figura A.3 do Anexo A), podendo fazer o acompanhamento deste último processo de mineração dos dados. É importante discorrer sobre esse fluxo, pois apesar de não fazer parte do LAD, é através dele que os modelos foram treinados para a predição do desempenho dos alunos e para gerar a análise dos mesmos, processo esse que está detalhado na metodologia deste trabalho e que serve de base para a próxima seção, que discorre a respeito do design do LAD.
2. **Modelos Salvos:** nessa tela, o usuário pode ver detalhes a respeito dos modelos que foram treinados previamente (Figura A.4 do Anexo A.).
3. **Filtros - LAD:** essa tela representa a visualização inicial da ferramenta desenvolvida, apresentando os filtros disponíveis para a geração da análise dos alunos (Figura 8).



**Figura 4.** Tela de Filtros - LAD do FMDEV [O autor, 2020]

É importante salientar que as telas descritas em 1. (Figura A.1, Figura A.2 e Figura A.3 do Anexo A) e 2. (Figura A.4 do Anexo A) pertencem a aplicação FMDEV, sendo a tela 3. (Figura 4) a que representa a visualização inicial da proposta de desenvolvimento do presente trabalho.

Além das três telas principais do FMDEV, o ambiente também dispõe de uma base de dados PostgreSQL, que é utilizada para salvar e armazenar dados referentes às ações dos usuários no framework, tais como, os dados de login dos usuários no FMDEV, informações sobre os treinamentos e modelos treinados, e os datasources (fontes de dados) cadastrados, por exemplo.

Durante o desenvolvimento do LAD, foi criada uma nova tabela na base de dados, cujo nome é “moodle”, com o propósito de armazenar os dados coletados da base de dados do NEAD - UPE no Moodle, dados estes que são logs dos alunos no próprio AVA Moodle, sendo organizados por colunas da tabela “moodle”, representando os indicadores utilizados no treinamento dos modelos de *machine learning*., por exemplo.

Em relação aos dados, eles foram disponibilizados no formato de .csv e importados na tabela “moodle” na base de dados do FMDEV.

## 4.2. Design do LAD

O LAD foi concebido para ser disponibilizado no ambiente *Web*, de modo a permitir aos usuários o seu fácil acesso. Além disso, durante a construção do LAD, a proposta foi para que o mesmo possa ser utilizado em *notebooks* e *desktops*, visto que, com telas maiores, é possível entregar ao usuário uma melhor experiência de visualização dos gráficos e resultados de análise à primeira vista (Few, 2013).

Para introdução ao design proposto, a Tabela 5 mostra as visualizações presentes no LAD.

**Tabela 5.** Visualizações implementadas no LAD

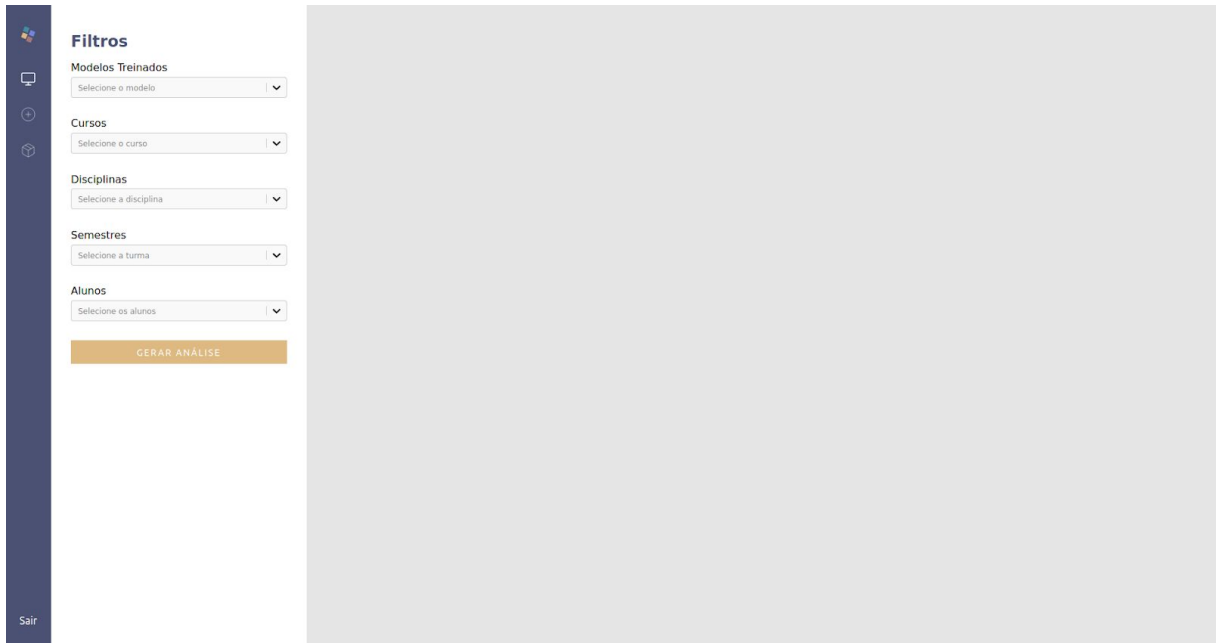
Visualização	Descrição
Desempenho satisfatório e insatisfatório	Informa o percentual de alunos cujo desempenho predito é satisfatório e o percentual de alunos cujo desempenho predito é insatisfatório.
Dicionário de indicadores	Apresenta quais os indicadores dos alunos utilizados no modelo de <i>machine learning</i> escolhido e o significado de cada um.
Análise preditiva do desempenho dos alunos	Apresenta a quantidade de alunos cujo desempenho predito é satisfatório e a quantidade de alunos cujo desempenho predito é insatisfatório.
Média dos indicadores dos alunos	Apresenta uma comparação da média dos valores de cada indicador dos alunos cujo desempenho predito é satisfatório com a média dos alunos cujo desempenho predito é insatisfatório.
Situação dos alunos	Permite a visualização das avaliações dos alunos, identificando-os por classe de desempenho predito, sendo as avaliações possíveis: nota do webquest, nota do fórum e nota da prova.



Indicadores do aluno e média da classe oposta	Apresenta um comparativo de cada indicador do aluno com a média de cada indicador da classe oposta ao desempenho predito para o aluno.
Análise descritiva do aluno	Apresenta um resumo das informações relativas às avaliações feitas pelo professor ao aluno.
Análise descritiva da turma	Apresenta um resumo das informações relativas às médias da turma das avaliações feitas pelo professor.

Na sequência, são apresentadas as visualizações do LAD e uma explicação detalhada do que cada visualização propõe.

- 1) A Figura 5 apresenta a tela inicial do LAD, onde são mostrados os filtros disponíveis para geração da análise de desempenho:
  - a) **Modelos treinados:** filtro que apresenta os modelos treinados no FMDEV e que estão disponíveis para escolha na geração da análise.
  - b) **Cursos:** filtro dos cursos disponíveis nas fontes de dados cadastradas no FMDEV.
  - c) **Disciplinas:** filtro das disciplinas disponíveis nas fontes de dados cadastradas no FMDEV, a partir da escolha do filtro de curso.
  - d) **Semestres:** filtro das turmas disponíveis nas fontes de dados cadastradas no FMDEV, a partir da escolha do filtro de disciplina.
  - e) **Alunos:** filtro dos alunos disponíveis nas fontes de dados cadastradas no FMDEV, a partir da escolha do filtro de semestre.



**Figura 5. Visualização inicial do LAD - Filtros** [O autor, 2020]

- 2) As Figuras 6 e 7 apresentam a tela de visão geral do LAD, onde são mostrados gráficos e informações relativos à análise geral da turma escolhida:
- a) **Cards de predição de desempenho:** blocos que mostram o percentual de alunos cujo desempenho predito é satisfatório e de alunos cujo desempenho predito é insatisfatório.
  - b) **Dicionário de indicadores:** tabela que apresenta as colunas “indicador” e “significado”, trazendo a descrição de cada indicador que foi utilizado no treinamento do modelo escolhido e que é utilizado nas visualizações.
  - c) **Análise preditiva de desempenho dos alunos:** gráfico de barras das duas classes do desempenho predito, satisfatório e insatisfatório, representando a quantidade de alunos em cada uma delas.
  - d) **Média dos indicadores dos alunos:** gráfico em barras, apresentado horizontalmente, que apresenta um comparativo da média de cada indicador das classes de desempenho predito, satisfatório e insatisfatório.

e) **Situação atual dos alunos:** gráfico de bolhas que permite a visualização das avaliações dos alunos, identificando-os por classe de desempenho predito, sendo as avaliações possíveis: nota do webquest, nota do fórum e nota da prova. Nesse gráfico, o eixo y é representado pela nota do webquest, o eixo x é representado pela nota do fórum e a largura da bolha representa a nota da prova. Ainda é possível visualizar, conforme a Figura 8, uma das possíveis variações do gráfico que representa a situação atual dos alunos, mostrando a sua dinamicidade, pois existem cursos que não utilizam a avaliação com webquest, cursos que não utilizam a avaliação do fórum, cursos que estão no início e, logo, só dispõem da primeira nota dos alunos, dentre outras possíveis situações no progresso do curso e que podem ser demonstradas por esse gráfico.

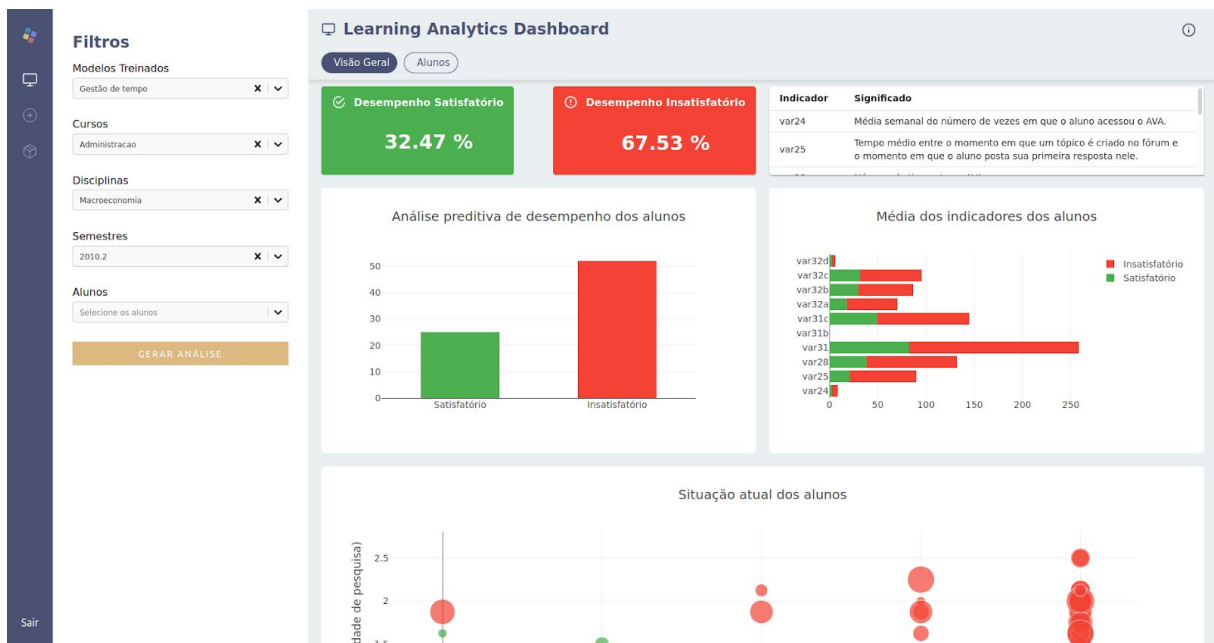


Figura 6. Tela de Visão Geral do LAD (1) [O autor, 2020]



Figura 7. Tela de Visão Geral do LAD (2) [O autor, 2020]

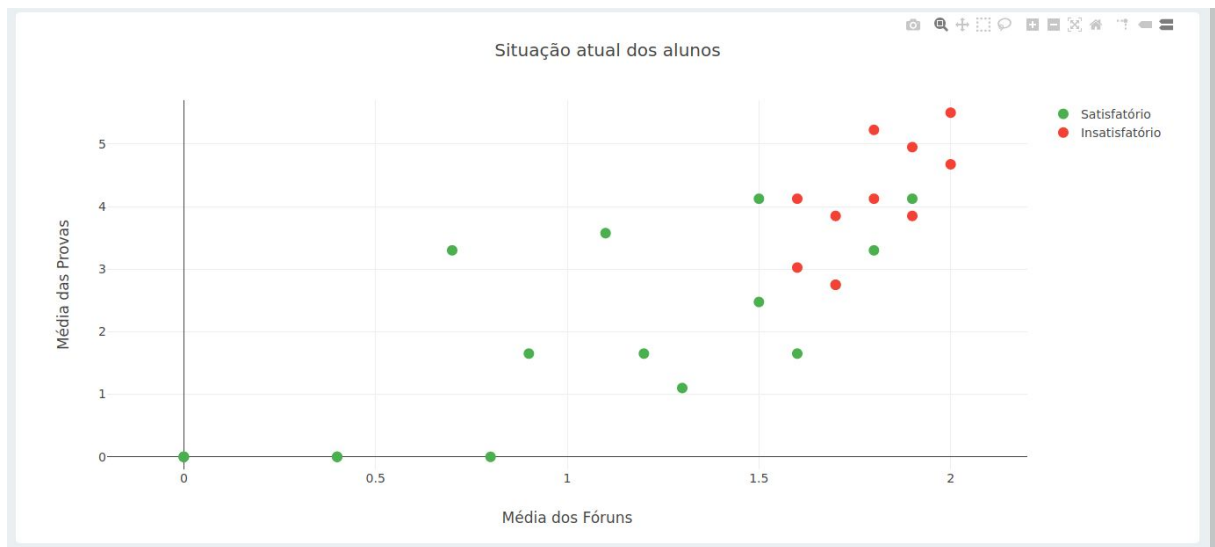


Figura 8. Variação do gráfico “Situação atual dos alunos” em uma turma onde não há avaliação de webquests [O autor, 2020]

3) As Figuras 9 e 10 apresentam a tela de Alunos do LAD, onde são mostrados gráficos e informações relativos a um aluno específico, selecionado por um filtro, e o comparativo dos indicadores desse aluno com médias da turma.

- a) **Indicadores do Aluno x Médias dos alunos da classe oposta:** gráfico de barras que compara o valor de cada indicador do aluno com a média de cada indicador dos alunos cuja classe predita é oposta a do aluno em questão. A Figura 9 mostra um cenário onde o aluno escolhido pelo filtro pertence a classe de desempenho satisfatório, enquanto que a Figura 10 mostra um cenário onde o aluno escolhido pelo filtro pertence a classe de desempenho insatisfatório.
- b) **Análise Preditiva:** card que mostra a quantidade de alunos da classe predita satisfatório e da classe predita insatisfatório, dando enfoque a qual classe pertence o aluno filtrado.
- c) **Análise descritiva do aluno:** card que mostra as informações relativas às avaliações do aluno filtrado.
- d) **Análise descritiva da turma:** card que mostra as informações relativas às médias das avaliações da turma, incluindo todos os alunos, não fazendo distinção por classe.

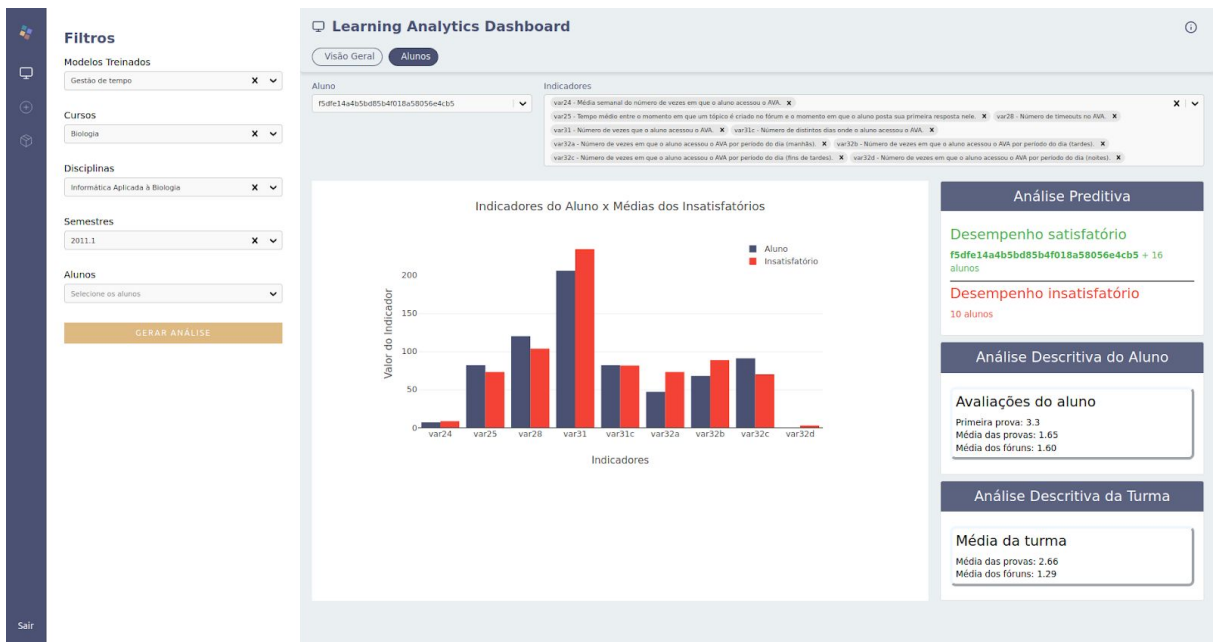
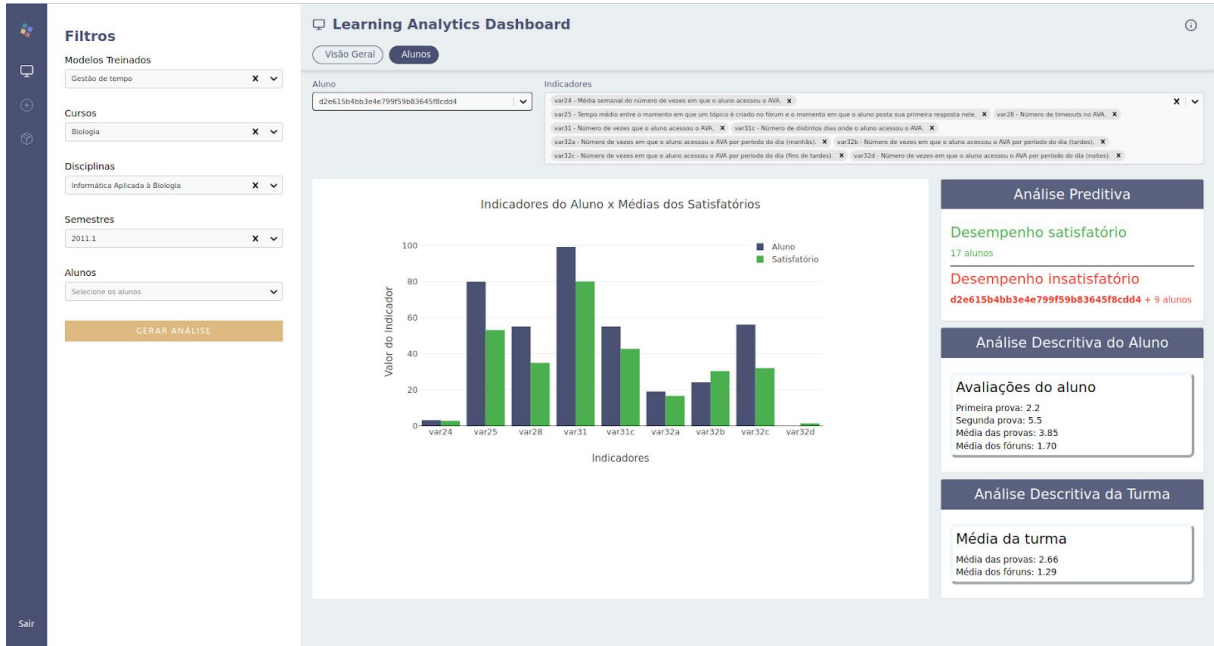


Figura 9. Tela de Alunos do LAD - aluno cujo desempenho predito é satisfatório [O autor, 2020]



**Figura 10. Tela de Alunos do LAD - aluno cujo desempenho predito é insatisfatório**  
[O autor, 2020]

# CAPÍTULO 5

## 5. VALIDAÇÃO E RESULTADOS

A validação do LAD foi realizada com professores do ensino superior que possuem experiência na modalidade de ensino a distância, onde foi feita uma análise da aceitação da ferramenta. Para a preparação do ambiente de validação, foram treinados modelos de *machine learning* no ambiente do FMDEV para serem utilizados na geração da análise feita pelo LAD. Além disso, foram disponibilizados, no LAD, dados de 2009 a 2016 da base de dados do NEAD - UPE no Moodle, referentes às disciplinas de Administração, Biologia, Letras e Pedagogia, para uso dos professores avaliadores na ferramenta.

Para que a validação fosse realizada, duas etapas foram percorridas:

1. Os professores receberam uma explicação inicial sobre o LAD, mostrando como acessá-lo e explicando que eles deveriam visualizar os dados, fazer uma análise sobre os tipos de visualizações e refletir sobre como essas informações visuais poderiam mudar seus conceitos e abordagem durante seu processo de ensino com suas turmas.
2. Após utilizarem o LAD pelo tempo que achassem necessário, os professores tiveram acesso a um questionário contendo dez perguntas, onde as mesmas estavam agrupadas em três seções: Utilidade percebida, Facilidade de uso percebida e Intenção de uso. Nesse questionário, os professores responderam a perguntas de modo a emitirem o seu grau de concordância ou discordância sobre algo, escolhendo um dentre os cinco seguintes valores da escala Likert: concordo plenamente, concordo, neutro, discordo e discordo plenamente. A Tabela 6 mostra as perguntas que compõem o questionário, baseando-se no modelo TAM, a seção a qual elas pertencem e a quantidade de professores que escolheram cada valor da escala Likert, relativo à pergunta.

**Tabela 6.** Questionário de validação do LAD respondido por professores

Seção	Pergunta	Quantidades de professores que escolheram cada valor na escala Likert
Utilidade percebida	P1. Usar o LAD em meu trabalho me permitiria realizar tarefas mais rapidamente.	Concordo plenamente: 1 Concordo: 5 Neutro: 0 Discordo: 1 Discordo plenamente: 0
	P2. Usar o LAD em meu trabalho aumentaria minha produtividade.	Concordo plenamente: 3 Concordo: 2 Neutro: 1 Discordo: 1 Discordo plenamente: 0
	P3. Usar o LAD aumentaria minha eficácia no trabalho.	Concordo plenamente: 2 Concordo: 4 Neutro: 0 Discordo: 1 Discordo plenamente: 0
	P4. Eu consideraria o LAD útil no meu trabalho.	Concordo plenamente: 1 Concordo: 4 Neutro: 1 Discordo: 1 Discordo plenamente: 0
Facilidade de uso percebida	P5. Aprender a operar o LAD seria fácil para mim.	Concordo plenamente: 3 Concordo: 3 Neutro: 1 Discordo: 0 Discordo plenamente: 0



	P6. Eu consideraria fácil fazer com que o LAD faça o que eu quero.	Concordo plenamente: 1 Concordo: 3 Neutro: 3 Discordo: 0 Discordo plenamente: 0
	P7. Minha interação com o LAD seria clara e compreensível.	Concordo plenamente: 1 Concordo: 3 Neutro: 3 Discordo: 0 Discordo plenamente: 0
	P8. Eu consideraria o LAD fácil de usar.	Concordo plenamente: 3 Concordo: 3 Neutro: 1 Discordo: 0 Discordo plenamente: 0
Intenção de uso	P9. Se o LAD estiver disponível, eu teria a intenção de usá-lo no meu trabalho com regularidade.	Concordo plenamente: 1 Concordo: 5 Neutro: 0 Discordo: 0 Discordo plenamente: 1
	P10. Se houver outras opções de visualização disponíveis no LAD, o mesmo seria a minha primeira opção.	Concordo plenamente: 2 Concordo: 3 Neutro: 2 Discordo: 0 Discordo plenamente: 0

A partir da análise dos resultados obtidos no questionário, pode-se perceber que: todas as perguntas da seção “Utilidade percebida” tiveram pelo menos 71,43% dos professores avaliando a utilidade da ferramenta como satisfatória (“concordo” ou

“concordo plenamente”); todas as perguntas da seção “Facilidade de uso percebida” tiveram pelo menos 57,15% dos professores avaliando a facilidade de utilização da ferramenta como satisfatória (“concordo” ou “concordo plenamente”). Vale ressaltar que 85,72% dos professores consideraram o LAD uma ferramenta fácil de usar, conforme a questão P8; todas as perguntas da seção “Intenção de uso” tiveram pelo menos 71,43% dos professores avaliando a intenção de uso da ferramenta satisfatoriamente (“concordo” ou “concordo plenamente”);

Não obstante os positivos resultados obtidos pela validação através da escala Likert, um dos professores fez um comentário geral a respeito do LAD: “A ferramenta é excelente para gestores da educação, mas não é própria para professores tomarem decisões.”, o que faz se refletir a respeito do usuário final do LAD, apesar de ser a opinião de um professor dentre os sete que participaram da validação.

Em suma, foi possível comprovar a utilidade, a facilidade e, conseqüentemente, a aceitação do LAD através de potenciais usuários. Além disso, através da questão “Se o LAD estiver disponível, eu teria a intenção de usá-lo no meu trabalho com regularidade.”, onde 85,72% dos professores escolheram ou valor “concordo plenamente” ou o valor “concordo”, comprovam que a utilização do LAD pode sim trazer benefícios no processo de ensino-aprendizagem no dia a dia dos professores e suas turmas.

# CAPÍTULO 6

## 6. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho foi proposta um *Learning Analytics Dashboard*, uma ferramenta para análise de dados a partir de modelos de mineração de dados educacionais, envolvendo tanto dados descritivos como preditivos, a partir de indicadores dos alunos que afetam o seu desempenho em uma disciplina, sendo estes provenientes de um ambiente virtual de aprendizagem.

Para validar a arquitetura proposta, foram apresentados resultados preliminares a respeito da aceitação do LAD e a intenção de utilizá-lo, a partir dos quais foi possível ver que a ferramenta desenvolvida pode trazer benefícios ao professor e a suas turmas, no dia a dia de lecionamento.

Apesar da validação feita, ela ainda é insuficiente, visto que apenas sete professores participaram do processo de validação do LAD. Por conta disso, como trabalho futuro, pretende-se considerar uma nova validação com mais professores envolvidos, podendo alcançar uma aceitação e intenção de uso de mais potenciais usuários da ferramenta. Além disso, pretende-se fazer um teste de usabilidade da ferramenta, de modo a receber o feedback dos usuários e desenvolver novas visualizações baseadas na experiência dos mesmos.

# BIBLIOGRAFIA

- [1] BAZIRE, M.; BRÉZILLON, P. **Understanding context before using it. In International and Interdisciplinary Conference on Modeling and Using Context** (p. 29-40). Springer, 2005.
- [2] BROWN, M. **Learning analytics: the coming third wave**. EDUCAUSE Learning initiative, 2011.
- [3] DAVIS, F. D. **Perceived usefulness, perceived ease of use and user acceptance of information technology**. *Management Information Systems Quarterly*, 13(3), 319-340, 1989.
- [4] DUVAL, Erik; VERBERT, Katrien. **Learning analytics, E-learning and Education**, 2014.
- [5] FARIA A. A., SALVADORI A. **Educação a Distância e seu movimento histórico no Brasil**. *Revistas das Faculdades Santa Cruz, Curitiba*, v. 8, n. 1, 2010.
- [6] FEW, S. **Information dashboard design: Displaying data for at-a-glance monitoring** (2nd ed.). Burlingame: Analytics Press.
- [7] FISHBEIN, M.; AJZEN, I. Belief. **Attitude, Intention and Behavior: An Introduction to Theory and Research**. Addison-Wesley, 1975.
- [8] GILFUS EDUCATION GROUP. **Learning Analytics**. Disponível em: <https://www.gilfuseducationgroup.com/learning-analytics/>. Acesso em: outubro de 2020.
- [9] KATS, Y. **Learning Management Systems and Instructional Design: Best Practices in Online Education**. Idea Group Inc (IGI), 2013.
- [10] LI, I.; DEY, A.; FORLIZZI, J. **A stage-based model of personal informatics systems**. In *Proceedings of CHI10: 28th International Conference on Human Factors in Computing Systems* (p. 557-566). New York, NY: ACM
- [11] LITTO, F.; FORMIGA, M. **Educação a Distância: o estado da arte**. São Paulo: Person Education do Brasil. 2 ed. 2012.

- [12] LIU, B. **Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data**. Berlin Heidelberg: Springer, 2006.
- [13] LONG, P.; SIEMENS, G. **Penetrating the fog: Analytics in learning and education**. p. 31-40, 2011.
- [14] PARK, Y.; JO, I.-H. **Development of the learning analytics dashboard to support students' learning performance**. p. 110-133, 2015.
- [15] PARK, Y.; JO, I.-H. **Factors that affect the success of learning analytics dashboards**. Association for Educational Communications and Technology, 2019.
- [16] PINTRICH, P. R. **The role of motivation in promoting and sustaining self-regulated learning**. International Journal of Educational Research, 1999.
- [17] RAMOS, D. K. **Perfil dos alunos de Licenciatura a Distância e aspectos que contribuem para aprendizagem**. Revista Reflexão e Ação, Santa Cruz do Sul, v. 21, n. esp., p. 199-220. Disponível em <https://online.unisc.br/seer/index.php/reflex>.
- [18] RAMOS, J. L. C.; RODRIGUES, R. L.; SILVA, J. C. S.; GOMES, A. S. **Analisando Fatores que Afetam o Desempenho de Estudantes Iniciantes em um Curso a Distância**. Anais do Simpósio Brasileiro de Informática na Educação. 2014, p. 99-108.
- [19] RABELLO, C. R. L. **Aprendizagem na Educação a Distância - dificuldades dos discentes na licenciatura em ciências biológicas na modalidade semipresencial**. Dissertação de Mestrado, 2007.
- [20] RODRIGUES, R. L.; RAMOS, J. L. C.; SILVA, J. C. S.; DOURADO, R. **Forecasting Students' Performance Through Self-Regulated Learning Behavioral Analysis**. International Journal of Distance Education Technologies. Volume 17, Issue 3. July-September. 2019.
- [21] ROSALES, G. **Modelo e framework para o desenvolvimento de ferramentas analíticas de apoio ao ensino, aprendizagem e gestão educacional**. Tese de Doutorado, 2014.

- [22] ROMERO, C.; VENTURA, S. **Educational data mining: A survey from 1995 to 2005**. Expert systems with applications, Elsevier, v. 33, n. 1, p. 135–146, 2007.
- [23] SILVA, R. G. **Desenvolvimento de um Ambiente de Aprendizado de Máquina Automatizado Integrável a Múltiplos AVAs**. Dissertação de Mestrado - Universidade de Pernambuco. Recife, p. 3, 2020.
- [24] ZALDIVAR, V. A. R. et al. **Monitoring Student Progress Using Virtual Appliances: A Case Study**. Computers & Education. 2012. 58 (4).
- [25] SHIH, T. K.; HUNG, J. C.; JIN, Q. **A survey of distance education challenges and technologies**. International journal of distance education technologies, 2003.
- [26] VERBERT, D. et al. **Learning analytics dashboard applications**. p. 3, 2013.
- [27] YIN, R. K. **Estudo de Caso: Planejamento e Métodos**. 5ª edição. Porto Alegre, Bookman, 2015.

# ANEXOS

## ANEXO A - TELAS DO FMDEV

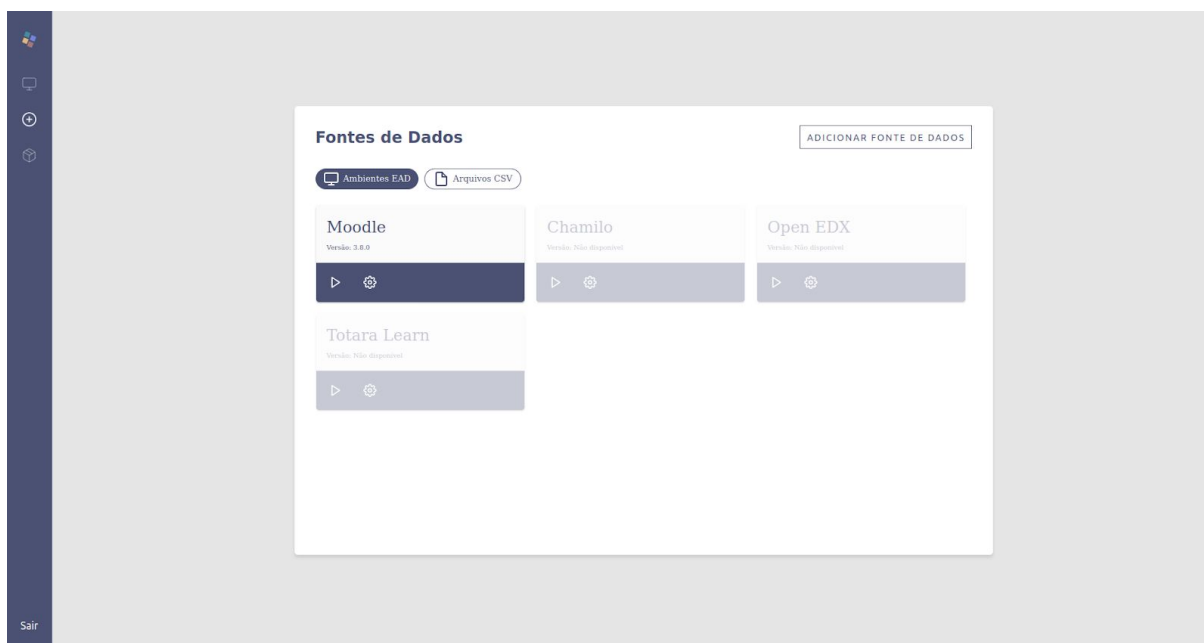


Figura A.1 Tela de Fonte de Dados do FMDEV [Silva, 2020]

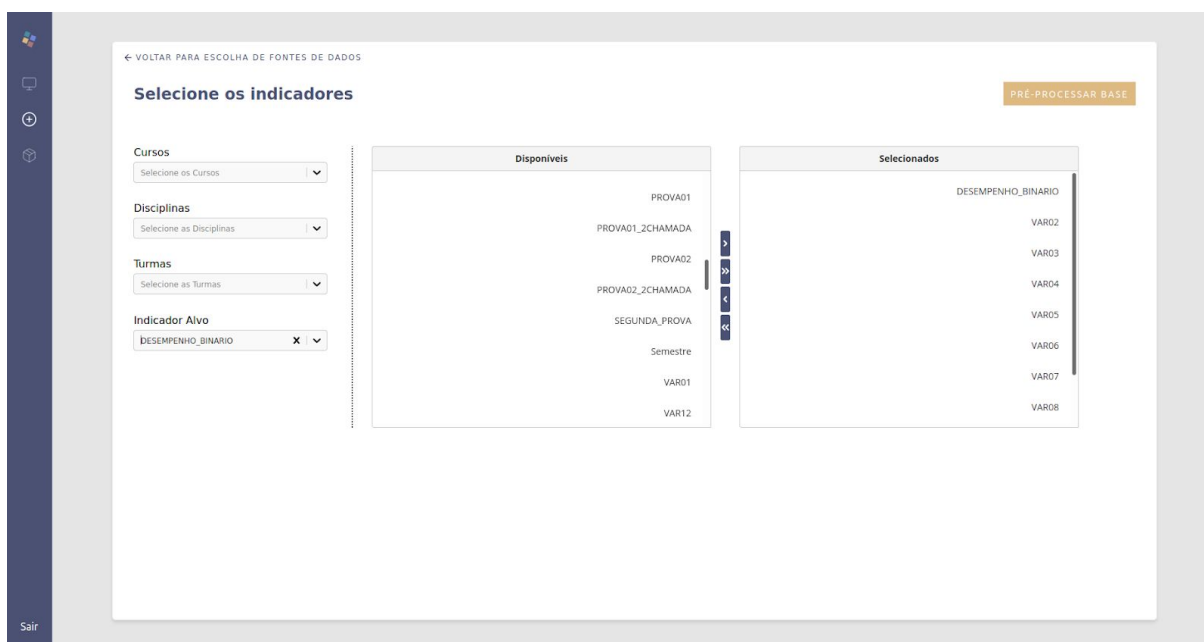


Figura A.2 Tela de Pré-processamento dos dados do FMDEV [Silva, 2020]

**Pré-processamento dos dados**

Fonte de dados: LMS/Moodle (Total de Instâncias : 30217)

CONFIGURAR TREINAMENTO

Indicador	Correlação	Faltante	Média	Desvio Padrão	Mínimo	Máximo	Ações
DESEMPENHO_BINARIO	Alvo	0	0.44	0.5	0	1	
VAR02	-0.17	0	0.19	0.89	0	19	
VAR03	-0.25	0	2.29	7.75	0	210	
VAR04	-0.41	0	13.83	33.58	0	627	
VAR05	-0.36	0	38.38	52.37	0	617	
VAR06	-0.09	0	0.03	0.22	0	8	
VAR07	-0.12	0	0.06	0.43	0	20	
VAR08	0.25	0	0.58	1.49	0	38	
VAR10	-0.28	0	1.1	2.57	0	41	
VAR09	-0.31	0	0.98	1.67	0	28	

**Parâmetros do treinamento**  
Use os parâmetros default ou altere-os caso seja necessário.

Percentual de treinamento:

Percentual de testes:

Qtd. máxima de treinos:

Qtd. de K-fold:

TREINAR CANCELAR

**Figura A.3** Tela de análise do Pré-processamento dos dados do FMDEV [Silva, 2020]

**Modelos Salvos**

Nome	Descrição	Criado em	Acurácia de teste	Última predição em	Predições realizadas	Ações
Gestão de tempo	Modelo baseado no constructo TM (Time management) para a seleção de variáveis.	28/10/2020 06:18:34	0.86	29/10/2020 02:27:57	14	...
Autoavaliação	Modelo baseado no constructo SA (Self-assessment) para a seleção de variáveis.	28/10/2020 05:29:27	0.71	29/10/2020 03:27:02	11	...
Definição de metas	Modelo baseado no constructo GS (Goal-setting) para a seleção de variáveis.	28/10/2020 05:51:28	0.86	29/10/2020 03:34:59	23	...
Estratégias de conclusão de tarefas	Modelo baseado no constructo TS (Task completion strategies) para a seleção de variáveis.	28/10/2020 02:38:54	0.84	29/10/2020 03:38:29	37	...
Em busca de ajuda	Modelo baseado no constructo HS (Help-seeking) para a seleção das variáveis.	28/10/2020 05:44:30	0.78	29/10/2020 05:35:50	16	...

**Figura A.4** Tela de Modelos Salvos do FMDEV [Silva, 2020]