



**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE ESTATÍSTICA E INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA APLICADA**

TARCÍSIO LOPES DE ALMEIDA SOUSA

**A TECNOLOGIA WEB SERVICES E SUA UTILIZAÇÃO COM
APLICAÇÕES DE DISPOSITIVOS MÓVEIS EM UMA
INSTITUIÇÃO DE ENSINO**

**RECIFE – PE
2023**

TARCISIO LOPES DE ALMEIDA SOUSA

**A TECNOLOGIA WEB SERVICES E SUA UTILIZAÇÃO COM
APLICAÇÕES DE DISPOSITIVOS MÓVEIS EM UMA
INSTITUIÇÃO DE ENSINO**

Dissertação submetida à Coordenação do programa de Pós-Graduação em Informática Aplicada da Universidade Federal Rural de Pernambuco, como parte dos requisitos necessários para obtenção do grau de Mestre.

ORIENTADOR: Prof. Dr. Moacyr Cunha Filho

**RECIFE – PE
2023**

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

S725t

Sousa, Tarcisio Lopes de Almeida

A tecnologia web services e sua utilização com aplicações de dispositivos móveis em uma instituição de ensino / Tarcisio Lopes de Almeida Sousa. - 2023.
78 f. : il.

Orientador: Moacyr Cunha Filho.
Inclui referências e apêndice(s).

Dissertação (Mestrado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Informática Aplicada, Recife, 2023.

1. Web services. 2. Dispositivos móveis. 3. Predição de desempenho acadêmico. 4. Computação em nuvem. 5. Avaliação de desempenho. I. Filho, Moacyr Cunha, orient. II. Título

CDD 004

TARCISIO LOPES DE ALMEIDA SOUSA

**A TECNOLOGIA WEB SERVICES E SUA UTILIZAÇÃO COM
APLICAÇÕES DE DISPOSITIVOS MÓVEIS EM UMA
INSTITUIÇÃO DE ENSINO**

Dissertação submetida à Coordenação do programa de Pós-Graduação em Informática Aplicada da Universidade Federal Rural de Pernambuco, como parte dos requisitos necessários para obtenção do grau de Mestre.

Aprovada em: 31/08/2023.

BANCA EXAMINADORA

Prof. Dr. Moacyr Cunha Filho (Orientador)
Universidade Federal Rural de Pernambuco
Departamento de Estatística e Informática

Profa. Dra. Silvana Bocanegra
Universidade Federal Rural de Pernambuco
Departamento de Estatística e Informática

Prof. Dr. Victor Casimiro Piscoya
Universidade Federal Rural de Pernambuco
Departamento de Tecnologia Rural

TARCISIO LOPES DE ALMEIDA SOUSA

Dedico esta dissertação aos meus pais e amigos, em especial àqueles que contribuíram para que este trabalho fosse realizado.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter andado comigo e nunca ter me abandonado. Mesmo diante das dificuldades encontradas, sempre me guiou pelo caminho correto, me dando forças para nunca pensar em desistir.

Agradeço ao Prof. Moacyr Cunha Filho pela orientação. Surgiu na hora certa para colocar o trem nos trilhos e desde então sou grato pelo acolhimento, incentivo e apoio que foram fundamentais para o desenvolvimento deste trabalho.

Agradeço também à minha mãe Evanilda Lopes de Almeida Sousa e ao meu pai Amano Pereira de Sousa, que deram os primeiros passos na minha educação, o que me motiva até hoje na busca constante por qualificação profissional.

Também meu muito obrigado aos colegas de turma, onde compartilhamos conhecimento por meses no laboratório de informática do DEINFO, sempre empenhados e dedicados aos estudos. Aos amigos de Petrolina-PE, em especial a Júnior Cavalcante, grande incentivador para que eu fizesse o Mestrado, me apoiou e fez contribuições valiosas para que eu chegasse até aqui. Também a Valério Leite, pelas dicas de programação e por estar disposto a ajudar.

Aos professores do PPGIA, em especial: Maria da Conceição Moraes Batista, Ricardo André de Souza, Ermeson Carneiro, Gilberto Cysneiros e Gustavo Callou; e também a todos os funcionários e colaboradores que sempre me atenderam muito bem, particularmente Eduardo Chaves, nosso secretário e estimado colega de turma. À UFRPE, pelo suporte e infraestrutura necessários para todas as atividades do curso.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. This study was financed in part by the Higher Education Personnel Improvement Coordination – Brazil (CAPES) – Finance Code 001.

Vocês foram muito importantes, meu eterno agradecimento.

RESUMO

Web service (serviço web) é um sistema de comunicação e interação entre diferentes aplicativos e plataformas pela internet, sendo uma forma de permitir que diferentes sistemas se comuniquem e compartilhem dados e funcionalidades, independentemente de suas tecnologias e linguagens de programação. Os web services são baseados em padrões e protocolos abertos, esses padrões permitem que os web services possam ser facilmente acessados e utilizados por diferentes aplicativos, sem qualquer restrição ao ambiente em que estejam sendo executados. Da mesma forma, aplicativos de dispositivos móveis utilizam web services para realizar a integração com sistemas distintos, como sistemas web. Um exemplo típico da utilização de web services no ambiente de computação móvel é na realização de operações em banco de dados, como consulta a dados e informações. Neste trabalho é apresentado o início do desenvolvimento de um protótipo de aplicativo para dispositivos móveis android, chamado de PredictApp, que pode auxiliar estudantes e professores na predição de desempenho acadêmico. Foi realizada uma revisão sistemática da literatura para levantamento dos principais fatores e tecnologias usadas na predição de desempenho acadêmico e como tais sistemas são passíveis de automação. Foi realizado ainda o estudo de técnicas de avaliação de desempenho do aplicativo e sua infraestrutura de comunicação com o banco de dados, utilizando algumas métricas de desempenho.

Palavras-chave: Web Services, Dispositivos Móveis, Predição de Desempenho Acadêmico, Computação em Nuvem, Avaliação de Desempenho.

ABSTRACT

Web service is a system of communication and interaction between different applications and platforms over the internet, serving as a way to enable different systems to communicate and share data and functionality, regardless of their technologies and programming languages. Web services are based on open standards and protocols, which allow them to be easily accessed and used by different applications, without any restrictions on the environment in which they are being executed. Likewise, mobile applications use web services to integrate with different systems, such as web systems. A typical example of the use of web services in the mobile computing environment is performing operations on databases, such as data queries and information retrieval. In this work, the initial development of a prototype mobile application for Android devices, called PredictApp, is presented. PredictApp aims to assist students and teachers in predicting academic performance. A systematic literature review was conducted to identify the main factors and technologies used in academic performance prediction and how such systems can be automated. Additionally, a study of performance evaluation techniques for the application and its communication infrastructure with the database was carried out, utilizing some performance metrics.

Keywords: Web Services, Mobile Devices, Academic Performance Prediction, Cloud Computing, Performance Evaluation.

LISTA DE FIGURAS

Figura 2.1 Processo de solicitação/resposta SOAP.....	19
Figura 2.2 Estrutura de uma mensagem SOAP.....	20
Figura 2.3 Exemplo de estrutura dos dados em arquivo JSON	21
Figura 2.4: Diferença entre os modelos de serviços em nuvem computacional	26
Figura 2.5: Classificação das técnicas de avaliação de desempenho	30
Figura 2.6: Definição de sucesso acadêmico	31
Figura 2.7: Algoritmos de classificação.....	32
Figura 3.1: Etapas do processo de pesquisa e seleção de artigos.....	40
Figura 4.1: Modelo Conceitual PredictApp	52
Figura 4.2: Processo de implantação do modelo.....	53
Figura 4.3: Tarefa Inserir Notas	54
Figura 4.4: Tarefa Inserir Faltas.....	54
Figura 4.5: Tarefa Consultar Notas e Faltas	55
Figura 4.6: Smartphone consumindo serviço.....	59
Figura 4.7: Casos de Uso	60
Figura 4.8: Diagrama de atividades	61
Figura 4.9: Diagrama de Pacotes	62
Figura 4.10: Diagrama de Classes.....	63
Figura 4.11: Diagrama de Sequência LogarNoApp.....	64
Figura 4.12: Diagrama de Sequência CarregarSpinner.....	65
Figura 4.13: Diagrama de Sequência ConsultarNotasFaltas.....	65
Figura 4.14: Tela login.....	66
Figura 4.15a: Tela Notas e Faltas.....	67
Figura 4.15b: Tela Spinner Semestres	67
Figura 4.16: Utilização de CPU	68
Figura 4.17: Utilização de Memória	69
Figura 4.18: Consumo de Energia.....	70

LISTA DE TABELAS

Tabela 1.1: Questões norteadoras	16
Tabela 2.1: Comparação de diferentes soluções de PaaS	27
Tabela 3.1: Questões norteadoras	36
Tabela 3.2: Resultado das pesquisas realizadas	37
Tabela 3.3: String de busca utilizada nas bases de dados	38
Tabela 3.4: Bases de dados da pesquisa e endereço eletrônico	39
Tabela 3.5: Critérios de Inclusão, Exclusão e Qualidade	40
Tabela 3.6: Artigos selecionados	41
Tabela 3.7: Variáveis significativas da empregabilidade.....	46
Tabela 3.8: Variáveis preditoras	46
Tabela 3.9: Variáveis relacionadas ao estudante	47
Tabela 3.10: Atributos e habilidades correspondentes.....	49
Tabela 3.11: Resultados dos artigos selecionados	50

SUMÁRIO

1 INTRODUÇÃO.....	12
1.1 DEFINIÇÃO DO PROBLEMA	15
1.2 MOTIVAÇÃO.....	15
1.3 DELIMITAÇÃO DO ESCOPO	15
1.4 QUESTÃO DE PESQUISA	16
1.5 OBJETIVO GERAL	16
1.5.1 Objetivos Específicos	16
1.6 ESTRUTURA DA DISSERTAÇÃO.....	17
2 FUNDAMENTAÇÃO TEÓRICA.....	18
2.1 TECNOLOGIAS PARA INTEGRAÇÃO DE DADOS	18
2.1.1 Web Services.....	18
2.1.2 Computação em Nuvem.....	25
2.1.3 Avaliação de Desempenho	28
2.2 PREDIÇÃO DE DESEMPENHO ACADÊMICO	30
2.2.1 Sucesso Acadêmico.....	30
2.2.2 Algoritmos Classificadores em Mineração de Dados Educacionais	32
2.2.3 Algoritmos Determinísticos.....	35
3 REVISÃO DA LITERATURA	36
3.1 PROCEDIMENTOS METODOLÓGICOS	36
3.2 ESTRATÉGIAS DE BUSCA.....	37
3.3 BASES DE DADOS E PROCESSO DE EXTRAÇÃO	37
3.4 CRITÉRIOS DE INCLUSÃO, EXCLUSÃO E QUALIDADE	38
3.5 RESULTADOS	40
3.5.1 Decision Support System to Enhance Students’ Employability using Data Mining Techniques for Higher Education Institutions (TAEZA-CRUZ; CAPILIKUMMER, 2023).....	42
3.5.2 Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling (GULERIA; SOOD, 2023)	42
3.5.3 Analysis and Prediction of Students’ Academic Performance and Employability Using Data Mining Techniques: A Research Travelogue (CRUZ; ENCARNACION, 2021)	43
3.5.4 Analyzing and Predicting Career Specialization using Classification Techniques (REPASO; CAPARIÑO, 2020)	44
3.5.5 Student and school performance across countries: A machine learning approach (MASCI et al., 2018).....	44
3.5.6 Respostas das questões de pesquisa.....	45
3.5.7 Considerações Finais do Capítulo	51

4 DESENHO DA ARQUITETURA	52
4.1 ETNOGRAFIA	53
4.1.1 Análise das Tarefas	53
4.2 ESPECIFICAÇÃO DE REQUISITOS	56
4.2.1 Requisitos Funcionais	56
4.2.2 Requisitos Não Funcionais	57
4.3 ARQUITETURA DA INTEGRAÇÃO	58
4.4 DIAGRAMAS	60
4.4.1 Diagrama de Casos de Uso	60
4.4.2 Diagrama de Atividades	61
4.4.3 Diagrama de Pacotes.....	61
4.4.4 Diagrama de Classes	62
4.4.5 Diagramas de Sequência.....	63
4.5 IMPLEMENTAÇÃO	66
4.6 AVALIAÇÃO DE DESEMPENHO	68
4.6.1 Utilização de CPU	68
4.6.2 Utilização de Memória.....	69
4.6.3 Consumo de Energia	70
5 CONCLUSÕES	71
REFERÊNCIAS	74
APÊNDICE A – FORMULÁRIO DE PERGUNTAS	78

1 INTRODUÇÃO

A Integração de dados é um dos campos de pesquisa mais antigos na área de Banco de Dados, tendo surgido logo após as empresas começarem a utilizar sistemas de banco de dados. Atualmente, é comum que empresas tenham sistemas de informação diferentes, mas coexistentes, para buscar oportunidades de negócios em mercados altamente competitivos. Neste cenário, a integração desses sistemas se tornou indispensável para o suprimento das necessidades dos negócios e dos clientes [Dittrich, 2004].

A rede mundial de computadores juntamente com a área dos negócios são os grandes responsáveis pelo crescente interesse na integração de dados. Esses dados, porém, dificilmente estarão em um formato comum. Combinar dados de fontes diferentes para obter uma resposta mais completa a uma consulta, ou para responder uma consulta complexa, não é uma tarefa simples. A mesma informação pode estar guardada com estruturas diferentes em bases diferentes, ou o mesmo dado pode conter valores contraditórios em fontes diversas.

O objetivo dos sistemas de integração de dados é fornecer uma visão integrada de diferentes fontes de dados, de forma que fique transparente para o usuário qual informação está sendo buscada em qual base. Isso significa que, apesar das bases serem distribuídas e heterogêneas, o usuário tem a impressão de utilizar um sistema centralizado e homogêneo [Amaral, 2005].

Integração de sistemas vem para auxiliar a integridade de dados e a produtividade de usuários, executando em *BackOffice*¹ tarefas que o usuário realizaria manualmente [Santos, 2007].

Web services são um conjunto de tecnologias utilizadas na integração de sistemas e na comunicação entre aplicações diferentes. Com esta solução é possível que aplicações possam interagir com outras, mesmo sendo desenvolvidas em plataformas diferentes. Os *Web Services* possuem recursos que permitem às aplicações enviar e receber dados em um formato comum. Apesar de terem sua própria linguagem, as aplicações podem traduzir as informações para uma

¹ É o núcleo de um sistema, que não é visto pelo usuário final. O *backoffice* possui tarefas específicas, coordena e gerencia atividades de um sistema, website etc. Qualquer problema que venha a ocorrer no *backoffice* prejudica todo um sistema, serviço, website, pois os arquivos, muitas vezes, são retirados dele e o *backoffice* precisa estar sempre em funcionamento para dar suporte aos usuários.

linguagem universal. Para as empresas, os *Web Services* podem trazer rapidez e eficiência nos processos. A comunicação entre aplicações passa a ser dinâmica e principalmente segura, pois não há intervenção humana [Mahmoud, 2005].

Uma aplicação Web oferece um serviço que exige a intervenção de um usuário, enquanto o *Web Service* facilita a interação direta entre as aplicações sem a necessidade de o usuário intervir [Andrade, 2008].

Atualmente, diversas instituições de ensino possuem um sistema acadêmico para uso de professores e alunos, além de funcionários do setor administrativo. Este sistema acadêmico é um sistema *web*, ou seja, com acesso através da internet, que permite aos alunos realizarem consulta ao boletim escolar, onde este é constituído de notas e frequências das disciplinas cursadas e também das disciplinas que o aluno esteja cursando.

Os professores de uma instituição de ensino possuem, dentre outras atribuições, a função de realizar o lançamento, via sistema acadêmico, das notas e frequências dos alunos que estão cursando sua disciplina. Normalmente, o professor realiza o lançamento das notas e frequências no sistema acadêmico da instituição de ensino, para que o aluno possa realizar a consulta ao boletim *on-line*.

Existe a necessidade de um sistema acadêmico e, conseqüentemente, de um aplicativo para dispositivos móveis que venha a ser utilizado como alternativa a essas consultas, mais especificamente, um aplicativo que possua requisitos para funcionar em *smartphones* e *tablets*, visto que estes equipamentos são os mais utilizados pelos alunos. Além disso, esse aplicativo poderá ser utilizado para realização de predições sobre o desempenho acadêmico dos alunos.

A educação passou por transformações significativas nos últimos anos devido aos avanços tecnológicos. A tecnologia desempenhou um papel fundamental na modernização do setor educacional, não apenas por meio do desenvolvimento de softwares educacionais, mas também na gestão eficiente de registros acadêmicos e na administração de instituições de ensino de forma virtual. Além disso, a internet proporcionou aos estudantes novas oportunidades de aprendizado, tornando o acesso ao conhecimento mais amplo e flexível. Essa combinação de elementos tecnológicos tem revolucionado a maneira como a educação é concebida, entregue e administrada [SOUZA; SANTOS, 2021].

Com o crescimento do volume e armazenamento de dados educacionais, juntamente com as informações dos alunos, torna-se imperativo utilizar recursos computacionais para otimizar a eficiência e qualidade na análise desses dados. Nesse contexto, a Mineração de Dados Educacionais (MDE) desempenha um papel fundamental, uma vez que essa metodologia pode ser aplicada em várias áreas e para diversas finalidades. Ela oferece uma abordagem poderosa para extrair *insights* valiosos dos dados educacionais, tornando os processos de ensino e aprendizado mais eficazes. A MDE permite a identificação de tendências, padrões e áreas de melhoria, auxiliando educadores e instituições de ensino na tomada de decisões embasadas em dados [VIVIAN et al., 2022].

A MDE envolve a aplicação de técnicas de mineração de dados, e em alguns casos, a aplicação de modelos probabilísticos em dados na área da educação. Para empregar a MDE de forma eficaz, é crucial compreender quais ferramentas são apropriadas e quais técnicas melhor se adequam ao contexto de estudo. Uma maneira de adquirir esse conhecimento é realizar uma Revisão Sistemática de Literatura (RSL), que consiste em uma análise dos materiais encontrados na literatura, com o objetivo de identificar e expor os métodos que podem ser replicados, proporcionando assim uma visão abrangente das ferramentas, técnicas e fatores explorados no contexto em estudo. Isso facilita a tomada de decisões e a implementação eficaz da MDE na educação [MARQUES et al., 2022].

Uma revisão sistemática da literatura (muitas vezes referida como revisão sistemática) é um meio de identificar, avaliar e interpretar todas as pesquisas disponíveis relevantes para uma determinada questão de pesquisa, ou área temática, ou fenômeno de interesse. Os estudos individuais que contribuem para uma revisão sistemática são chamados de estudos primários; uma revisão sistemática é uma forma de estudo secundário [KEELE et al., 2007].

O método MDE tem como objetivo primordial na área acadêmica a predição do desempenho dos alunos ao longo de seu curso. Isso permite que as instituições de ensino realizem intervenções quando necessário, visando aprimorar o processo de aprendizagem dos alunos. Essa investigação é de grande importância não apenas para a vida acadêmica e profissional dos estudantes, mas também para a própria universidade. A qualidade do trabalho da instituição é diretamente refletida no desempenho de seus alunos, o que torna essas análises essenciais para a melhoria constante do ensino e para a avaliação da qualidade educacional [SOUZA; SANTOS, 2021].

Diante desse contexto, este trabalho também realizará uma revisão sistemática da literatura, de modo que se possa verificar os trabalhos realizados sobre predição de desempenho acadêmico, na busca de identificar as principais variáveis envolvidas, os modelos e tecnologias utilizadas e a possibilidade de automação através do uso de um aplicativo para dispositivos móveis.

1.1 DEFINIÇÃO DO PROBLEMA

O sucesso dos estudantes possui um papel fundamental nas instituições de ensino, pois frequentemente serve como um indicador do desempenho da própria instituição. A identificação precoce de alunos em situação de risco, juntamente com a implementação de medidas preventivas, pode ter um impacto significativo em seu êxito acadêmico. Técnicas de aprendizado de máquina têm sido amplamente adotadas para fins de predição. No entanto, a aplicação eficaz e eficiente de métodos de mineração de dados envolve muitas decisões, que vão desde como definir o sucesso do estudante, passando por quais atributos do aluno devem ser enfatizados, qual o método de aprendizado de máquina é mais apropriado para o problema em questão, se a solução é passível de automação e quais as tecnologias podem ser utilizadas em sua implementação.

1.2 MOTIVAÇÃO

De acordo com a definição do problema apresentado, a motivação para esta pesquisa reside na necessidade de investigar quais são as técnicas de mineração de dados educacionais, com um foco especial nas determinísticas, que tenham demonstrado melhor acurácia nos resultados e se é passível de automação um sistema de avaliação e predição de desempenho de estudantes com previsibilidade de desempenho profissional. Essa investigação tem o propósito de obter as informações necessárias para dar início ao desenvolvimento de um aplicativo com a funcionalidade de exibir os resultados para professores e estudantes.

1.3 DELIMITAÇÃO DO ESCOPO

Embora o tema predição de desempenho acadêmico seja um dos objetos de estudo deste trabalho, onde inclusive é realizada uma revisão sistemática da literatura, não é objetivo deste trabalho realizar a implementação dos sistemas de predição.

1.4 QUESTÃO DE PESQUISA

Para abordar as questões apresentadas na Revisão Sistemática da Literatura, foram formuladas três questões que desempenham o papel de guiar o estudo e a análise dos artigos encontrados, conforme a Tabela 1.1:

Tabela 1.1: Questões norteadoras

Questões	Descrição
Q1:	É passível de automação um sistema de avaliação e predição de desempenho de estudantes deterministicamente com previsibilidade de desempenho profissional?
Q2:	Quais tecnologias atuais poderiam ser usadas em um protótipo nesse sentido?
Q3:	Ranks (como TOEFL, GRE, ENEM, PISA ²) são eficazes na elaboração de diretrizes e guias para as instituições de ensino terem melhores egressos?

1.5 OBJETIVO GERAL

Este trabalho pretende dar uma contribuição no *review* da área de avaliação de desempenho acadêmico com o auxílio de ferramentas sintáticas e iniciar um protótipo de um aplicativo que fará uso de Web Services para automação deste processo.

1.5.1 Objetivos Específicos

- i. Compreender as tecnologias utilizadas na integração de dados;
- ii. Compreender o estado da arte sobre predição de desempenho acadêmico;
- iii. Através de uma revisão sistemática, identificar as técnicas de mineração de dados educacionais com melhor acurácia;

² Programa de Avaliação Internacional de Alunos é uma pesquisa internacional trienal (iniciada no ano 2000) que tem como objetivo avaliar os sistemas de educação em todo o mundo, testando as habilidades e conhecimentos de estudantes de 15 anos de idade.

- iv. Iniciar o desenvolvimento de um protótipo de aplicativo para dispositivos móveis que exiba o boletim escolar e o resultado da predição de desempenho acadêmico de estudantes.

1.6 ESTRUTURA DA DISSERTAÇÃO

O trabalho está organizado em 5 capítulos correlacionados. O Capítulo 1, Introdução, apresentou por meio de sua contextualização o tema proposto neste trabalho. Da mesma forma foram estabelecidos os resultados esperados por meio da definição de seus objetivos e apresentadas as limitações do trabalho permitindo uma visão clara do escopo proposto.

O Capítulo 2 apresenta a fundamentação teórica, onde é realizada uma revisão da literatura sobre os temas Tecnologias para Integração de Dados e Predição de Desempenho Acadêmico.

O Capítulo 3 apresenta a realização de uma revisão sistemática da literatura sobre o tema predição de desempenho acadêmico. São apresentados os procedimentos metodológicos, estratégias de busca, as bases de dados utilizadas na pesquisa, o processo de extração e os critérios de inclusão, exclusão e de qualidade. São apresentados os resultados, com uma síntese dos 5 trabalhos escolhidos pelos critérios e a descrição das respostas para as questões de pesquisa.

O Capítulo 4 apresenta o desenho da arquitetura, com a especificação dos requisitos, artefatos gerados nas fases de análise e projeto, bem como a sua implementação e a avaliação de desempenho através da técnica de medição.

No Capítulo 5 são tecidas as conclusões do trabalho, relacionando os objetivos identificados inicialmente com os resultados alcançados. São ainda propostas possibilidades de continuação da pesquisa desenvolvida a partir das experiências adquiridas com a execução do trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta uma revisão da literatura para um melhor entendimento deste trabalho. O referencial teórico representa uma importante etapa de pesquisa, proporcionando uma estrutura para estabelecer a importância do estudo e também estabelecer uma referência para comparar resultados [Creswell, 2010]. Inicialmente, são apresentados os conceitos sobre web services. Esses conceitos abrangem os tipos de arquitetura de web services e uma abordagem sobre o uso dessa tecnologia em dispositivos móveis. Em seguida, são apresentados os conceitos básicos sobre computação em nuvem, a classificação dos modelos de serviços e a plataforma em nuvem do Heroku, adotada nesse trabalho. Logo após, são apresentados os conceitos básicos de avaliação de desempenho e suas classificações. Em seguida, são apresentados conceitos relacionados à predição de desempenho acadêmico.

2.1 TECNOLOGIAS PARA INTEGRAÇÃO DE DADOS

2.1.1 Web Services

Existem tecnologias como HTTP e HTML que são utilizadas pelos browsers para exibir páginas da internet. Porém, estas tecnologias não são bem sucedidas na integração de aplicações. Com o uso de web services, possibilita-se que diferentes aplicações comuniquem entre si e utilizem recursos diferentes [Mahmoud, 2005].

A abordagem de web services é baseada em um conjunto de normas padronizadas que são amplamente aceitas e utilizadas. Esta aceitação generalizada torna possível para os clientes e serviços se comunicarem e entenderem uns aos outros através de uma ampla variedade de plataformas e linguagens [Ort, 2005].

Construção de web services é um dos pilares mais importantes no mundo do desenvolvimento de sistemas. Web services são utilizados como forma de integração e comunicação de sistemas, de modo que um sistema possa realizar uma chamada para um serviço de outro sistema a fim de obter informações. Estas chamadas podem enviar e receber informações em diversos formatos, sendo que atualmente os mais populares são XML e JSON [Lecheta, 2015].

Embora exista um padrão para acesso aos serviços do web service, construí-lo não deixa o desenvolvedor preso a uma linguagem de programação. Por exemplo, é possível desenvolver um web service com a linguagem PHP, mas consumir seus serviços em qualquer outra linguagem, como Java, C#, Python, etc. Existem diversas maneiras de construir web services, mas as mais conhecidas são web services em SOAP ou REST.

Web Services SOAP

O SOAP (*Simple Object Access Protocol*) é um protocolo de comunicação de web services, que consiste em uma mensagem no formato XML utilizada como entrada e saída de web services do tipo SOAP, descritos por uma WSDL (*Web Services Description Language*), que é um grande arquivo XML que descreve todos os métodos disponíveis neste serviço, ou seja, os métodos que podem ser chamados e executados [Lecheta, 2015].

Como é baseado em texto e autodescritivo, as mensagens SOAP podem transmitir informações entre serviços em ambientes de computação heterogêneos sem se preocupar sobre problemas de conversão. Além do WSDL, existe outra especificação que é a UDDI (*Universal Description, Discovery and Integration*). WSDL define um padrão para descrever métodos de um Web Service, e UDDI define regras baseadas em XML para publicar informações de Web Service. As mensagens são trocadas usando GET/POST sobre HTTP através do protocolo SOAP, independentemente de onde o cliente esteja na rede [Ort, 2005]. A Figura 2.1 exibe como é realizado o processo de solicitação/resposta utilizando o protocolo SOAP:



Figura 2.1 Processo de solicitação/resposta SOAP
Fonte: Meniya; Jethva, 2012.

Na prática, o SOAP é um grande arquivo XML trafegando entre sistemas para fazer a comunicação. Pode-se observar a estrutura de uma mensagem SOAP na Figura 2.2:

```

POST /ExemploHelloWS/endpoint HTTP/1.1
Host: www.server.com
Content-Type: text/xml; charset="utf-8"
Content-Length: 322
SOAPAction: ""

<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:ns1="http://hello">

  <soapenv:Body>
    <ns1:sayHello>
      <ns1:name>friend</ns1:name>
    </ns1:sayHello>
  </soapenv:Body>
</soapenv:Envelope>

HTTP/1.1 200 OK
Content-Type: text/xml; charset="utf-8"
Content-Length: 367

<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:ns1="http://hello">

  <soapenv:Body>
    <ns1:sayHelloResponse>
      <ns1:return>Hello friend!</ns1:return>
    </ns1:sayHelloResponse>
  </soapenv:Body>
</soapenv:Envelope>

```

Figura 2.2 Estrutura de uma mensagem SOAP

Fonte: ULISBOA, 2018.

A primeira parte da Figura 2.2 mostra a mensagem SOAP de entrada, e a segunda parte, a mensagem SOAP de saída. Isso significa que, se algum sistema precisar se comunicar com este web service, ele precisará fazer uma requisição nesta página passando este XML do protocolo SOAP como parâmetro. A resposta também será um grande arquivo SOAP no formato XML.

Pelo motivo do SOAP ser um grande arquivo, ele começou a perder espaço para web services RESTful, pois estes apresentam uma sintaxe mais enxuta e podem enviar e receber informações em formatos mais leves. As aplicações precisam economizar recursos ao trafegar

informações, principalmente no desenvolvimento *mobile*, e justamente por isso o SOAP está a cada dia que passa perdendo mais espaço para o REST.

Web Services RESTful

O REST (*Representational State Transfer*) é um estilo de arquitetura de software utilizado principalmente para criar serviços web e auxiliar na integração de sistemas [Lecheta, 2015].

Cada vez mais utilizado no mundo inteiro, o REST é uma alternativa que surgiu para o desenvolvimento de web services, além do já conhecido SOAP. No REST é utilizado o protocolo HTTP para criar serviços que retornam dados geralmente nos formatos XML ou JSON.

Além do já citado formato XML, outro formato bastante utilizado por web services REST é o formato JSON. A Figura 2.3 mostra um exemplo de arquivo de saída com informações no formato JSON:

```
{
  "CurrentWeather": {
    "Location": "Curitiba Aeroporto , Brazil (SBCT) 25-315 049-10W 908M",
    "Time": "Aug 02, 2015 - 11:00 AM EDT / 2015.08.02 1500 UTC",
    "Wind": "from the NW (310 degrees) at 13 MPH (11 KT):0",
    "Visibility": "greater than 7 mile(s):0",
    "Temperature": "73 F (23 C)",
    "DewPoint": "42 F (6 C)",
    "RelativeHumidity": "33%",
    "Pressure": "30.15 in. Hg (1021 hPa)",
    "Status": "Success"
  }
}
```

Figura 2.3 Exemplo de estrutura dos dados em arquivo JSON

Fonte: Lecheta, 2015.

Comparando-se ao XML, o JSON possui uma estrutura menor, pois nele não é necessário abrir e fechar as tags, e é mais enxuto, diminuindo o tráfego de informações na rede, o que é primordial em aplicativos para dispositivos móveis. De forma prática, o JSON não passa de um objeto ou uma estrutura de dados representada em JavaScript.

Segundo Lecheta [2015], o REST possui características próprias, como por exemplo:

- Cada requisição ao serviço deve retornar de forma independente e não deve possuir estado. Para atingir este objetivo, pode-se utilizar o protocolo HTTP, o qual já é amplamente adotado na internet.
- Utiliza uma URI (*Uniform Resource Identifier*), ou seja, uma sintaxe universal para identificar recursos. No REST, cada recurso deve conter uma URI para que seja possível consultar informações sobre ele.
- Utilização de tipos de conteúdo (*mime-type*) para solicitar conteúdo e retornar conteúdo. Esse recurso permite que o cliente do serviço especifique se deseja que o conteúdo seja retornado em XML ou JSON, por exemplo.
- Possui um conjunto de operações padronizadas, baseadas nos métodos do protocolo HTTP, conforme a seguir:
 - GET utilizado para consultar;
 - POST utilizado para inserir;
 - PUT utilizado para atualizar;
 - DELETE utilizado para deletar.

O termo RESTful foi originado para designar que um determinado serviço segue os conceitos, ou em termos gerais, os princípios básicos do REST. Então se um web service for construído com base nas características do REST, pode-se dizer que este é um web service RESTful.

Web Services no Contexto de Dispositivos Móveis

A tecnologia Web Services se tornou um padrão da indústria para conectar recursos remotos e de bases distintas e, junto a isso, os dispositivos móveis se tornaram uma parte essencial do dia-a-dia das pessoas. As pessoas usam dispositivos móveis a qualquer hora e em qualquer lugar, podem usar seus celulares para verificar e-mails ou abrir outros aplicativos que usam a internet. Web Services admitem acessibilidade abrangente, permitindo a mobilidade do usuário, uma vez que supera as restrições de localização física da computação convencional. No entanto, a computação

móvel também requer uma tecnologia que conecte sistemas móveis a um ambiente de computação distribuída convencional [Meniya; Jethva, 2012].

A integração da computação móvel com a tecnologia Web Services trouxe muitas vantagens para ambos os lados. Os dispositivos móveis se tornaram computacionalmente úteis, de modo que podem ser tanto clientes como provedores de serviços na web.

A aplicação de modelos de comunicação de Web Services, como SOAP, à computação móvel, pode resultar em sobrecargas de desempenho inaceitáveis. Este problema vem de dois fatores: Primeiro, a codificação e decodificação de mensagens SOAP são detalhadas e baseadas em XML, o que acaba consumindo bastantes recursos e com isso os usuários podem sofrer com o baixo desempenho; Em segundo lugar, a diferença de desempenho e qualidade entre a comunicação cabeada e a sem fio deve ser considerada, bem como restrições no ambiente móvel como, por exemplo, velocidade do processador, duração limitada da bateria e conexão lenta e oscilante.

REST é uma arquitetura de software modelada de acordo com a forma como os dados são representados, acessados e modificados na web. Na arquitetura REST, os dados e a funcionalidade são considerados recursos e esses recursos são acessados por meio de *Uniform Resource Identifiers* (URIs), normalmente links na web. Os recursos são acionados por meio de um conjunto de operações simples e bem definidas. A arquitetura REST é fundamentalmente uma arquitetura cliente-servidor e é projetada para usar um protocolo de comunicação sem estado, normalmente HTTP. Na arquitetura REST, clientes e servidores trocam representações de recursos usando interface e protocolo padronizados. Esses princípios incentivam os aplicativos REST a serem simples, leves e com alto desempenho [Fielding, 2000].

Web Services RESTful expõe recursos (dados e funcionalidades) por meio de URIs da Web e usam os quatro principais métodos HTTP para criar, recuperar, atualizar e excluir recursos. Web Services RESTful mapeiam os quatro métodos HTTP para as chamadas ações CRUD: criar, recuperar, atualizar e excluir.

Web Services REST visam ser simples, e isso é feito limitando os tipos de operações que podem ser executadas em um recurso. Os fundadores do REST afirmam que [Fielding, 2000]:

- Fornece tempos de resposta aprimorados e características de carregamento do servidor devido ao suporte para cache.

- Melhora a escalabilidade do servidor reduzindo a necessidade de manter o estado de comunicação.
- Requer menos do software do lado do cliente para ser escrito do que outras abordagens, porque um único navegador pode acessar qualquer aplicação e qualquer recurso.
- Depende menos do software do fornecedor do que dos mecanismos que colocam estruturas de mensagens adicionais sobre o HTTP.
- Fornece funcionalidade equivalente quando comparada a abordagens alternativas de comunicação.
- Não requer um mecanismo separado de descoberta de recursos, devido ao uso de hiperlinks no conteúdo.
- A capacidade de tipos de documento, como HTML, de evoluir sem quebrar a compatibilidade com versões anteriores ou futuras.
- A capacidade dos recursos de adicionar suporte para novos tipos de conteúdo à medida que são definidos, sem descartar ou reduzir o suporte para tipos de conteúdo mais antigos.

As vantagens do uso de web services RESTful sobre web services convencionais para dispositivos móveis incluem menor tamanho de mensagem e tempo de resposta. Os resultados da comparação de desempenho entre SOAP convencional e RESTful mostram o óbvio alto desempenho RESTful sobre SOAP. A hospedagem de web services RESTful no ambiente de nuvem também melhorará o desempenho. Portanto, RESTful oferece uma solução perfeitamente boa para a maioria das implementações, com maior flexibilidade e menor *overhead* com computação em nuvem móvel [Meniya; Jethva, 2012].

Meniya e Jethva [2012], em suas avaliações comparativas incluíram operações de concatenação de strings, adição de números inteiros e adição de números flutuantes, onde os resultados da avaliação de desempenho mostraram as vantagens do uso de web services RESTful no ambiente de computação móvel. Estes autores consideram que os benefícios da utilização de web services RESTful no desenvolvimento para dispositivos móveis é superior ao uso de web services SOAP.

2.1.2 Computação em Nuvem

A computação em nuvem é um serviço de processamento baseado na web, por meio do qual recursos compartilhados, software e informações são fornecidos aos usuários pela internet [Meniya; Jethva, 2012]. Nos dias atuais, a computação em nuvem também tem impacto na indústria móvel. O aumento da conectividade de rede e a popularidade dos *smartphones* e *tablets* aumentaram consideravelmente o uso de aplicativos em dispositivos móveis. Na Computação em Nuvem Móvel, o armazenamento e processamento de dados ocorrem fora do dispositivo móvel e os resultados são exibidos na “palma da mão”. Uma de suas principais vantagens é que os usuários não estão mais conectados a um computador de mesa em casa ou no escritório, pois podem acessar todas as funcionalidades de um aplicativo que utiliza computação em nuvem de qualquer local, desde que tenham uma conexão com a internet.

O *data center* é o núcleo de processamento e armazenamento na nuvem. A virtualização é a tecnologia central do *data center*, pode-se dizer que a virtualização é uma das principais tecnologias da computação em nuvem. Graças a esta tecnologia, a abstração de recursos físicos pode ser realizada de forma que vários servidores virtuais podem ser criados como um único servidor físico ou como vários servidores físicos. O servidor virtual possui características similares ao servidor físico, isso pode otimizar os recursos de computação, de modo que os recursos do servidor físico possam ser particionados para fornecer serviços para vários clientes [Peixoto, 2012].

A computação em nuvem oferece, então, vantagens como compartilhamento de serviços, sistemas operacionais e aplicativos em rede, facilitando o gerenciamento dos modelos de negócios sem se preocupar com detalhes da infraestrutura [Machida et al., 2011]. O armazenamento em nuvem é um modelo de armazenamento *on-line*, no qual os dados são armazenados em *pools*³ de armazenamento virtualizados, geralmente hospedados por terceiros. As empresas de hospedagem operam *data centers* e as pessoas que precisam que seus dados sejam hospedados alugam capacidade de armazenamento deles. Os operadores de *data center*, virtualizam recursos de acordo com os requisitos do cliente e os expõem como *pools* de armazenamento, que os próprios clientes podem usar para armazenar dados. Fisicamente, o recurso pode abranger vários servidores. As

³ Um pool de armazenamento é uma coleção de discos físicos. Um pool de armazenamento permite a agregação de armazenamento e expansão da capacidade.

plataformas de nuvem oferecem além de armazenamento, poder computacional e diversos outros recursos na internet.

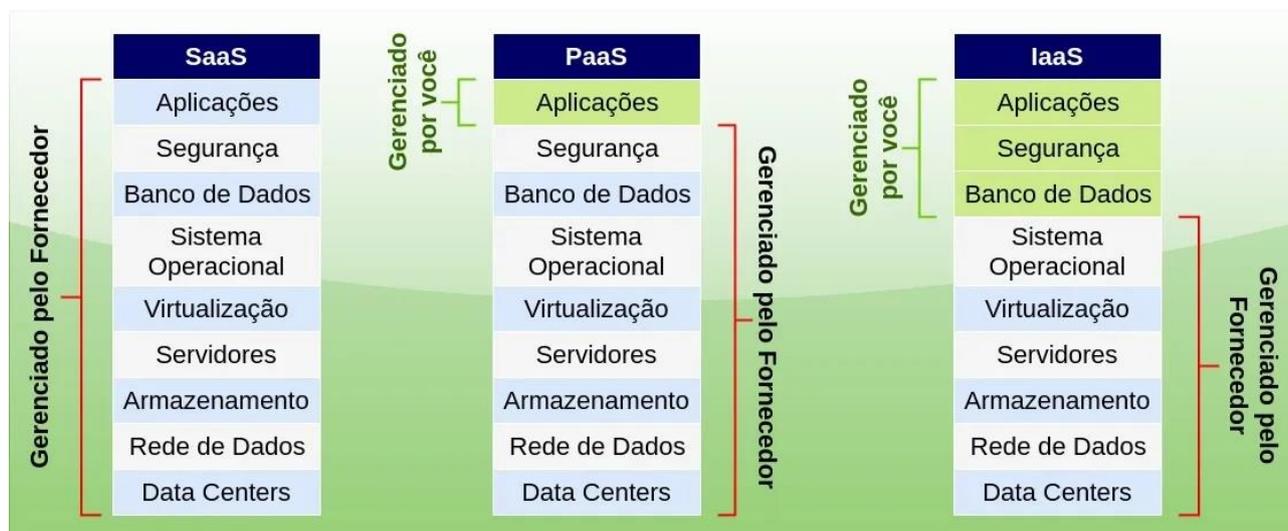


Figura 2.4: Diferença entre os modelos de serviços em nuvem computacional
Fonte: Costa et al., 2017.

Os três modelos de serviços em nuvem computacional podem ser classificados de acordo com o seu modelo de negócio (ZOTA; PETRE, 2014):

- **Infraestrutura como Serviço (Infrastructure as a Service - IaaS)**: É um modelo de serviço que fornece recursos de computação como hardware, armazenamento e rede.
- **Plataforma como Serviço (Platform as a Service - PaaS)**: É um modelo de serviço que o usuário pode utilizar para o desenvolvimento e execução de aplicações.
- **Software como Serviço (Software as a Service - SaaS)**: É um modelo de serviço onde são fornecidos aplicativos executados em uma infraestrutura de nuvem.

A grande disponibilidade de recursos e serviços, além do crescente acesso de usuários à computação em nuvem, vem favorecendo o aumento acelerado de hospedagem em diversas soluções de TI. A Figura 2.4 mostra algumas diferenças destes três modelos.

A disponibilidade dos serviços dos web services RESTful torna-se muito melhor após armazená-los para o ambiente em nuvem no modelo PaaS, pois nesse modelo o gerenciamento dos web services pode ser realizado pelo desenvolvedor. Assim sendo, foi utilizado o ambiente de hospedagem em nuvem gratuito fornecido pelo Heroku [Heroku, 2023]. Na seção 2.5.1, são

apresentadas mais informações sobre a plataforma de computação em nuvem do Heroku, que é o ambiente adotado neste trabalho pelo seu amplo nível de utilização.

Heroku

Heroku é uma plataforma como serviço (PaaS) que executa aplicativos em contêineres⁴ virtuais que são executados em um ambiente de tempo de execução confiável. Heroku chama esses contêineres de Dynos. O Heroku permite que o desenvolvedor dimensione o aplicativo instantaneamente apenas aumentando o número de dinômetros ou alterando o tipo de dinômetro em que o aplicativo é executado [Danielsson et al., 2013].

Foram consideradas várias soluções de PaaS disponíveis com base nas características mencionadas na Tabela 2.1 para selecionar uma plataforma adequada para a hospedagem dos web services e do banco de dados do protótipo do aplicativo. As plataformas PaaS consideradas incluíram Amazon Elastic Beanstalk, Dokku, Google App Engine, Heroku e OpenShift. Na Tabela 2.1, '+' é considerado a melhor opção, '-' é a pior opção e '0' é neutro.

Tabela 2.1: Comparação de diferentes soluções de PaaS

Características	Amazon Elastic Beanstalk	Dokku	Google App Engine	Heroku	OpenShift
Rapidez na implantação	+	+	+	+	+
Redução de complexidade	+	-	+	+	+
Extensibilidade	-	-	-	+	+
Inovação	+	+	+	+	+
Formas de pagamento	-	-	-	-	0
Facilidade para integração	+	+	+	+	+
Portabilidade	+	+	+	+	+
Capacidade	0	+	+	+	-
Custos	+	+	-	+	0
Usuários por aplicativo	0	-	+	+	-

Fonte: Dados da pesquisa

Após a comparação, o Heroku acabou sendo a melhor opção. Portanto, foi escolhido o Heroku para hospedagem do web service e banco de dados porquê ele oferece escalabilidade,

⁴ É um software usado para empacotar e isolar virtualmente aplicativos para permitir maior escalabilidade, disponibilidade e portabilidade em diversos ambientes de computação, incluindo sistemas em nuvem, máquinas virtuais (VMs) e alguns sistemas operacionais.

diminuição da complexidade, segurança e um menor custo. Heroku contém todas as dependências necessárias para a aplicação e é executado em um host compartilhado. Sempre que necessário, o número de dinâmômetros de uma aplicação executados podem ser ampliados ou reduzidos. Em vez de gastar tempo implementando o armazenamento de dados, pode-se simplesmente escolher um dos armazenamentos de dados disponíveis.

2.1.3 Avaliação de Desempenho

A avaliação de desempenho de sistemas de computação consiste em um conjunto de técnicas classificadas em medição e modelagem. A técnica de modelagem pode ser subclassificada ainda em analítica e simulação [LILJA, 2005; MENASCÉ; ALMEIDA, 2005].

A medição do sistema é o método mais direto de avaliação de desempenho e envolve necessariamente o monitoramento do sistema durante o período em que este está sendo acionado por uma carga de trabalho. Esta carga deve ser selecionada com cuidado para se obter resultados representativos. As técnicas de medição, embora forneçam respostas precisas, nem sempre podem ser utilizadas ou são muito caras, por exemplo, se o sistema estiver em ambiente de produção. Além disso, embora a carga de trabalho real seja uma excelente alternativa para uma representação fiel do sistema, às vezes esta opção não é aplicável. Pode ocorrer, por exemplo, se o tamanho da carga não corresponder aos dados coletados por motivo da influência de perturbações externas. Para facilitar a vida dos analistas de projeto que realizam medições, alguns mecanismos são utilizados para gerar carga de trabalho [LILJA, 2005].

Escolher uma carga de trabalho é tão importante quanto definir a estratégia de avaliação de desempenho que deve ser seguida. As diferentes estratégias são baseadas no conceito de evento, que é uma mudança no estado do sistema. A definição exata de evento depende da métrica que está sendo abordada. Por exemplo, um evento pode ser definido como um acesso ao disco, uma referência da memória ou uma operação de comunicação da rede de computadores [LILJA, 2005; MENASCÉ; ALMEIDA, 2005].

As ferramentas desenvolvidas para avaliar o desempenho de sistemas computacionais alteram o comportamento do que está sendo medido. Quanto maior a quantidade de dados e resolução fornecidos pela ferramenta de medição, maior a perturbação inserida por esta ferramenta.

Essa perturbação causada pela ferramenta de medição acaba retornando dados coletados menos confiáveis [LILJA, 2005; MENASCÉ; ALMEIDA, 2005].

A qualidade da medição pode ser indicada pela exatidão e precisão do instrumento medição. A exatidão é a diferença absoluta entre o valor medido e o valor de referência. E a precisão é a menor variação entre os valores obtidos por várias medições de uma característica especial do sistema. Medições muito precisas são muito melhor agrupadas em torno de um valor medido. Medições imprecisas tendem a uma maior variação. A precisão dessas medições é indicada por sua variação em torno da média. E exatidão é a diferença entre o valor da média e o valor medido [LILJA, 2005; MENASCÉ; ALMEIDA, 2005].

A simulação é usada tanto para avaliação de desempenho quanto para validação de modelos analíticos. Ao contrário das medições, as simulações são baseadas em modelos abstratos do sistema, então não requer a implantação total do sistema. Assim, os modelos utilizados durante a simulação foram desenvolvidos por abstração das principais características do sistema, bem como sua complexidade e nível de abstração varia de acordo com o sistema. É possível até controlar de forma mais eficaz os valores assumidos por parâmetros do sistema. Isso facilita a obtenção de informações em termos de avaliação de desempenho [LILJA, 2005; MENASCÉ; ALMEIDA, 2005].

A técnica analítica usa fórmulas fechadas ou um sistema de equações para descrever o comportamento do sistema. Métricas de interesse podem ser fornecidas através da solução de fórmulas fechadas ou da solução exata ou aproximada de um conjunto de equações fornecidas por algoritmos matemáticos numéricos [BOLCH, 2006].

Os fatores que influenciam e interferem no comportamento do sistema são modelados e representados por meio de parâmetros de equações matemáticas. Embora esses modelos aceitem parâmetros específicos do sistema, podem ser adaptados para outros sistemas. Ao construir modelos, deve-se levar em consideração sua complexidade e praticidade. Os modelos permitem uma análise ampla e minuciosa dos efeitos causados pelos parâmetros definidos nas equações sobre a aplicação. Além disso, também podem ser estabelecidas possíveis relações entre cada parâmetro considerado. Para confirmar os resultados alcançados pelos modelos desenvolvidos, a técnica analítica pode compará-los com os valores reais medidos em testes experimentais. Esses valores devem comprovar as predições realizadas através de modelos. Esta técnica apresenta custos mais

baixos em comparação com outras técnicas de avaliação de desempenho [LILJA, 2005; MENASCÉ; ALMEIDA, 2005; BOLCH, 2006].

Dependendo das restrições de tempo e recursos, pode ser possível usar mais de uma tecnologia que permita uma melhor compreensão do sistema e forneça informações mais detalhadas para ajudar na tomada de decisões. Com o auxílio de ferramentas de modelagem, simulação e medição, obtêm-se informações detalhadas sobre o sistema. Usando a técnica de medição, resultados mais precisos podem ser obtidos porque é a técnica que melhor corresponde ao ambiente real analisado. No entanto, a modelagem oferece vários métodos alternativos de avaliação de desempenho, que podem ser obtidos através de simulação ou modelagem analítica, onde os modelos analíticos lidam com métodos probabilísticos, teoria de filas, modelos de Markov ou redes de Petri [Callou et al., 2011]. A Figura 2.5 mostra que a análise de desempenho pode ser feita de várias maneiras diferentes:

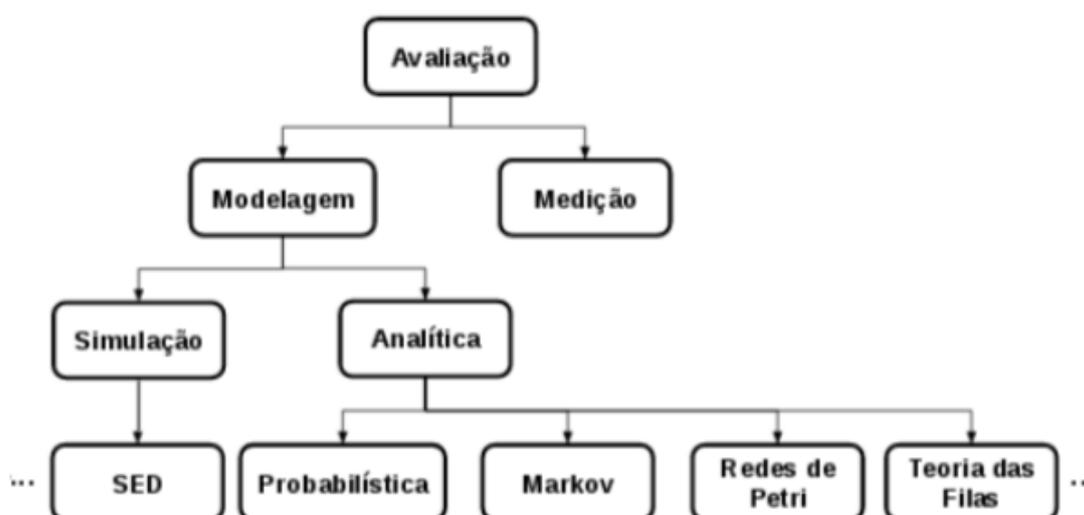


Figura 2.5: Classificação das técnicas de avaliação de desempenho
Fonte: Callou et al., 2011.

2.2 PREDIÇÃO DE DESEMPENHO ACADÊMICO

2.2.1 Sucesso Acadêmico

O sucesso do aluno é caracterizado pelo seu desempenho acadêmico, participação ativa em atividades de valor educacional, satisfação pessoal, aquisição de conhecimentos desejados, desenvolvimento de habilidades e competências, continuidade nos estudos, realização de metas educacionais e desempenho após a conclusão do curso [KUH et al., 2006].

O sucesso dos estudantes é um elemento de extrema importância nas instituições de ensino superior, uma vez que é considerado um critério fundamental para a avaliação da qualidade dessas instituições.

Embora existam diversas definições na literatura, seis componentes são considerados mais importantes para o sucesso acadêmico: Desempenho acadêmico, satisfação, aquisição de aptidões e competências, persistência, cumprimento de objetivos de aprendizagem e sucesso profissional, conforme a Figura 2.6 [YORK et al., 2015]:

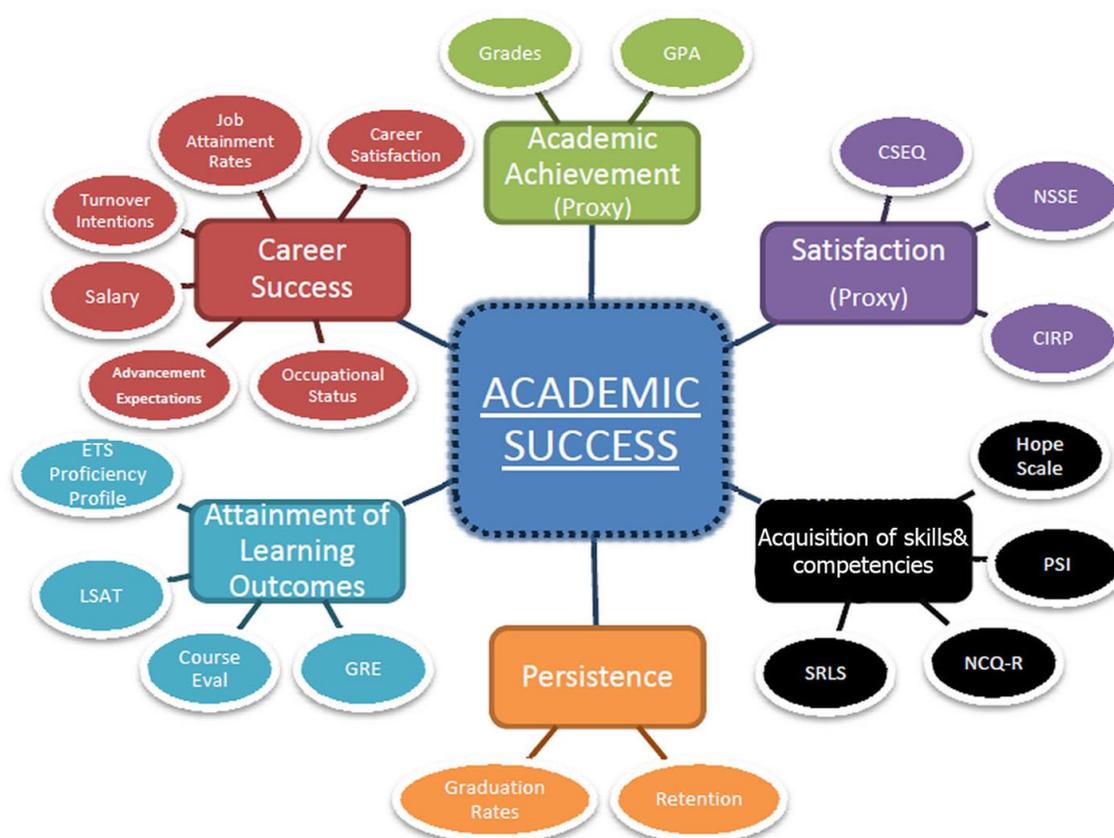


Figura 2.6: Definição de sucesso acadêmico
Fonte: YORK et al., 2015.

A orientação profissional correta impulsiona o desempenho dos alunos e aumenta seu nível de motivação. Os sistemas baseados em orientação profissional são essenciais para avaliar as habilidades dos alunos e, conseqüentemente, os recrutadores atribuem o cargo certo a eles [VIHERÄKOSKI, 2020]. A orientação profissional é necessária para solucionar os problemas que surgem na mente dos pais em relação aos seus filhos, desempenhando um papel importante ao guiar os pais na escolha do curso certo para seus filhos [YANG et al., 2020].

2.2.2 Algoritmos Classificadores em Mineração de Dados Educacionais

A classificação é um dos métodos mais comumente usados para categorizar grandes conjuntos de dados. Essa abordagem envolve adaptação de dados e algoritmos de aprendizado supervisionado. O objetivo é estabelecer a relação entre a variável de interesse de forma qualitativa e outras variáveis observadas [HUSSAIN et al., 2018].

A análise de classificação pode ser usada para consultar, determinar ou prever o comportamento por meio de um algoritmo. Funciona gerando um conjunto de dados de treinamento contendo várias características essenciais e resultados potenciais. A tarefa do algoritmo de classificação é descobrir como o conjunto de características chega ao seu final, o que determina e atribui categorias a um grupo de dados a ser avaliado com mais precisão. Essa abordagem utiliza métodos matemáticos como árvores de decisão, J48, Floresta Aleatória, SVM, programação linear, redes neurais e estatísticas [JASSIM, 2021]. A Figura 2.7 apresenta os algoritmos de classificação em mineração de dados educacionais.

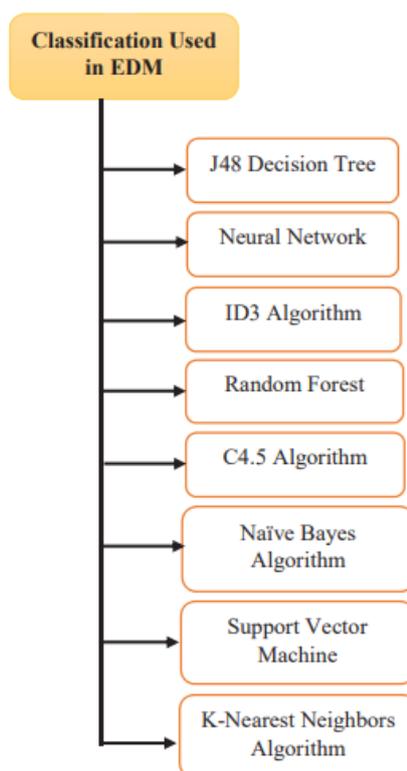


Figura 2.7: Algoritmos de classificação
Fonte: Adaptado de JASSIM, 2021.

A classificação é uma tecnologia de mineração de dados que atribui classes a diferentes conjuntos de dados para apoiar previsões e análises. A classificação é uma das maneiras de processar eficientemente grandes conjuntos de dados. O objetivo da classificação é prever corretamente a população alvo nos dados para cada evento. A técnica de classificação envolve a aprendizagem e a classificação dos dados. É o método de mineração de dados mais amplamente utilizado e é usado para desenvolver classes e atribuir conjuntos de dados a cada uma delas. Em resumo, se a variável alvo for discreta, então isso é um problema de classificação, e se a variável alvo for contínua, é uma tarefa de regressão [JASSIM, 2021].

Árvore de Decisão J48

O algoritmo J48 é usado para classificar várias aplicações e obter resultados de classificação corretos. O J48 é uma extensão do ID3 [LEE et al., 2020]. Recursos adicionais do J48 são suas características ausentes e a escolha da poda da árvore; é um algoritmo para criar uma árvore de decisão gerada pelo C4.5 (uma extensão do ID3) [KAUR et al., 2020]. Também é conhecido como um classificador estatístico. Para classificar a árvore de decisão, é necessário um banco de dados.

Rede Neural

Uma rede neural é uma coleção de algoritmos que têm como objetivo identificar relacionamentos fundamentais em vários dados por meio de uma estrutura que imita o funcionamento do cérebro humano [QIU et al., 2020]. Nesse contexto, as redes neurais se referem a sistemas neurais artificiais. Redes neurais podem ser adaptadas a entradas em evolução. A estrutura, assim, alcança o melhor resultado sem a necessidade de redesenhar os parâmetros de saída [HU; RANGWALA, 2020]. A ideia de redes neurais, que tem suas raízes em inteligência artificial, tornou-se muito comum no desenvolvimento de sistemas de negociação, como bolsa de valores.

Algoritmo ID3

O algoritmo ID3 foi desenvolvido por Ross Quinlan [AHMED et al., 2020] e é usado para gerar uma árvore de decisão a partir de um conjunto de dados. O ID3 é uma introdução ao C4.5 e é comumente usado para processar aprendizado de máquina e linguagens naturais. A árvore resultante é usada para identificar possíveis amostras.

Floresta Aleatória

A Floresta Aleatória é uma teoria de aprendizado estatístico e o método é usado com múltiplas árvores de decisão para fazer previsões e usar a votação para obter os resultados finais da previsão [HUNG et al., 2020]. O tamanho do subconjunto é sempre o mesmo que o tamanho da amostra de entrada original, mas a substituição é usada para retirar exemplos.

Algoritmo C4.5

O algoritmo C4.5 é usado para mineração de dados como classificação de árvore de decisão para tomar decisões com base em um conjunto de dados fornecido (previsões univariadas ou multivariadas) [MAHAJAN; SAINI, 2020]. O termo "classificador" aqui se refere a um método de mineração de dados que leva os dados que precisamos para identificar a nova categoria de dados e prever essa categoria.

Algoritmo de Naïve Bayes

Um grupo de algoritmos de classificação baseados no teorema de Bayes são os classificadores Naïve Bayes. Não é um único algoritmo; é um grupo de algoritmos. Todos eles compartilham um princípio comum: cada par de características é tratado como independente. O Naïve Bayes é fornecido com um conjunto de dados de treinamento padronizado [MAHESHWARI et al., 2020]. Portanto, é tratado como um algoritmo de aprendizado supervisionado. Ele ajuda a prever a probabilidade de ocorrência com base nas circunstâncias que são conhecidas.

Máquina de Vetores de Suporte

A Máquina de Vetores de Suporte (SVM) é um algoritmo de aprendizado de máquina supervisionado para classificação, regressão e detecção de variações. O algoritmo SVM tem como objetivo encontrar uma separação entre dois grupos de objetos, presumindo que quanto maior a separação, mais robusta a classificação [WAHEED et al., 2020]. A Máquina de Vetores de Suporte (SVM) opera de forma semelhante ao algoritmo C4.5, exceto que a SVM não utiliza árvores de decisão. Para classificar o conhecimento em dois grupos, a SVM aprende conjuntos de dados e determina o hiperlink.

Algoritmo dos K Vizinhos Mais Próximos

Os K Vizinhos Mais Próximos (KNN) é um algoritmo simples que armazena todos os estados existentes e classifica novos casos com base em uma escala semelhante. No início da

década de 1970, o KNN já havia sido usado como um método não paramétrico em estimação estatística e reconhecimento de padrões [YULIANTO et al, 2020]. O KNN é um algoritmo de aprendizado supervisionado e classificação de padrões que nos ajuda a determinar a classe à qual a nova entrada pertence (valor de teste) ao selecionar os K Vizinhos Mais Próximos e calcular a distância entre eles.

2.2.3 Algoritmos Determinísticos

Algoritmos determinísticos são procedimentos ou conjuntos de regras que, quando seguidos, produzem um resultado previsível e consistente. Isso significa que, dadas as mesmas entradas iniciais e as mesmas condições iniciais, um algoritmo determinístico sempre produzirá o mesmo resultado [ARAUJO et al., 2001].

Os algoritmos determinísticos são usados em uma ampla variedade de aplicações, desde cálculos matemáticos simples até processos complexos em ciência da computação, engenharia, física e muitos outros campos. Eles são contrastados com algoritmos não determinísticos, nos quais as saídas podem variar mesmo com as mesmas entradas devido a fatores aleatórios ou probabilísticos. A previsibilidade e consistência dos algoritmos determinísticos são muitas vezes uma vantagem em situações em que a precisão e a repetibilidade são essenciais, como em sistemas de controle, criptografia e em muitos aspectos da computação [REZENDE, 2003].

Algoritmos de mineração de dados educacionais podem ser tanto determinísticos quanto não determinísticos, dependendo da abordagem e da técnica utilizada. Alguns algoritmos de mineração de dados educacionais, como árvores de decisão, regressão logística e muitos algoritmos estatísticos, são determinísticos. Eles produzem resultados previsíveis e consistentes com base nas entradas e dados de treinamento fornecidos [ARAÚJO et al., 2019].

No entanto, também existem abordagens de mineração de dados educacionais que podem incorporar elementos não determinísticos. Por exemplo, o uso de técnicas de aprendizado de máquina, como redes neurais, pode envolver aspectos estocásticos durante o treinamento, tornando o processo de aprendizado não estritamente determinístico. Além disso, a aleatoriedade em alguns modelos pode tornar o processo de mineração de dados educacionais menos determinístico [ARAÚJO et al., 2019].

3 REVISÃO DA LITERATURA

Os trabalhos apresentados nesta seção são o resultado de uma revisão sistemática da literatura de publicações relevantes dos últimos 5 anos (2018-2023), com o propósito de identificar pesquisas semelhantes a esta e explorar o que está sendo pesquisado sobre predição de desempenho acadêmico no intuito de responder às questões de pesquisa, relacionadas na Tabela 3.1.

Tabela 3.1: Questões norteadoras

Questões	Descrição
Q1:	É passível de automação um sistema de avaliação e predição de desempenho de estudantes deterministicamente com previsibilidade de desempenho profissional?
Q2:	Quais tecnologias atuais poderiam ser usadas em um protótipo nesse sentido?
Q3:	Ranks (como TOEFL, GRE, ENEM, PISA) são eficazes na elaboração de diretrizes e guias para as instituições de ensino terem melhores egressos?

3.1 PROCEDIMENTOS METODOLÓGICOS

A revisão sistemática da literatura (RSL) adere a uma abordagem metodológica claramente delineada, demonstrando contribuições para um tema ou fenômeno de maneira imparcial e passível de repetição, enquanto examina perguntas específicas de pesquisa [KITCHENHAM, 2012].

A RSL possibilita reunir e avaliar a comprovação empírica de uma área de estudo específica por meio da análise de resultados de pesquisas pertinentes ao tema de interesse, permitindo, dessa forma, tirar conclusões acerca das perguntas de pesquisa previamente definidas. Para atingir esse propósito, emprega-se mecanismos de busca com o intuito de elaborar uma revisão abrangente da pesquisa relacionada à mineração de dados educacionais [NASCIMENTO et al., 2023].

Neste trabalho, a revisão sistemática da literatura será desenvolvida de acordo com o modelo proposto por Kitchenham (2012), abrangendo as seguintes fases: planejamento, implementação e resultados.

A etapa inicial compreendeu a elaboração do plano para esta pesquisa, estabelecendo o objetivo principal desta revisão sistemática: compreender quais mídias são mais enfatizadas pela mineração de dados educacionais, bem como os métodos utilizados para coletar e analisar esses dados, e de que forma a mineração de dados contribui para aprimorar a gestão educacional, possibilitando uma tomada de decisão mais eficaz.

Na etapa subsequente, a segunda fase, foram formuladas três questões-chave destinadas a orientar a pesquisa e a interpretação dos resultados. Isso foi realizado com o objetivo de fornecer uma contextualização qualitativa abrangente do tema e, ao mesmo tempo, permitir a análise quantitativa de alguns dos dados da pesquisa relacionados a essa temática.

3.2 ESTRATÉGIAS DE BUSCA

A estratégia de busca desta revisão foi dividida em duas etapas distintas. A primeira etapa envolveu a definição da *string* de busca, conforme Tabela 3.2, com palavras-chave relevantes para o estudo. A segunda etapa, por sua vez, consistiu na pesquisa em bibliotecas digitais, que servirão como fontes de estudo do tema proposto [KITCHENHAM, 2012].

Tabela 3.2: String de busca utilizada nas bases de dados

Operadores	OR	AND
<i>String de busca</i>	"academic success" OR "academic achievement" OR "student success" OR "predicting professional performance" OR "career success" OR "Employability" AND "Educational data" AND "syntactic tools" OR "classification techniques" OR "classifier algorithms" AND "automation" OR "prototype" OR "web application" OR "mobile app"	

3.3 BASES DE DADOS E PROCESSO DE EXTRAÇÃO

O processo de busca de artigos foi conduzido utilizando as principais bases de dados eletrônicas e científicas que fossem relevantes para abordar as questões Q1, Q2 e Q3. Um trabalho acadêmico relevante é um documento científico que atende a critérios como, conteúdo acadêmico de revistas científicas, teses, dissertações, conferências, relatórios técnicos e outros tipos de publicações.

São trabalhos que possuem qualidade e credibilidade, pois são revisados por pares (*peer-reviewed*) em caso de artigos de revistas, ou avaliados e aprovados por um comitê acadêmico em caso de teses e dissertações. São trabalhos disponíveis online e acessíveis ao público, incluindo publicações em repositórios institucionais, sites de universidades, periódicos online e outras fontes de acesso público. A relevância de um trabalho acadêmico também pode ser influenciada pelo número de citações que ele recebe de outros trabalhos acadêmicos. Trabalhos que são amplamente citados tendem a ser considerados mais relevantes.

As bases de dados utilizadas neste trabalho foram do Portal de Periódicos da CAPES e Google Acadêmico (*Scholar Google*), conforme Tabela 3.3. Estas bases indexam conteúdo acadêmico de várias fontes na web, incluindo artigos, teses, dissertações, relatórios técnicos, repositórios institucionais e outras fontes que estão disponíveis *on-line*. Muitos dos artigos que aparecem nos resultados de pesquisa dessas bases de dados podem estar disponíveis em bases de dados como Scopus, Web of Science, IEEE Xplore, e outras.

Tabela 3.3: Bases de dados da pesquisa e endereço eletrônico

Base de Dados	Endereço Eletrônico	Tipo de Busca
Periódicos CAPES	https://www.gov.br/capes/pt-br	Automática
Scholar Google	https://scholar.google.com.br/	Automática

Fonte: Dados da pesquisa

A Revisão Sistemática da Literatura empregou uma abordagem de busca automática, levando em consideração critérios de inclusão, exclusão e qualidade dos artigos encontrados. Os artigos científicos considerados foram limitados a uma janela temporal de até 5 anos de publicação.

3.4 CRITÉRIOS DE INCLUSÃO, EXCLUSÃO E QUALIDADE

Um passo crucial na condução de uma Revisão Sistemática da Literatura é estabelecer critérios que ajudem na seleção adequada do material bibliográfico. Esses critérios geralmente são organizados em três categorias: Critérios de Inclusão, Critérios de Exclusão e Critérios de Qualidade. Os artigos foram obtidos por meio da *string* de busca que foi aplicada nas bases de

dados, utilizando as questões formuladas para a pesquisa. Em seguida, eles foram selecionados de acordo com os critérios estabelecidos, conforme descritos na Tabela 3.4.

Tabela 3.4: Critérios de Inclusão, Exclusão e Qualidade

Critérios de Inclusão (I)
(I.1) Artigos de periódicos, dissertações e teses de repositórios institucionais;
(I.2) Publicações que estejam dentro do tema proposto, independente de idioma;
(I.3) Documentos disponíveis publicamente para leitura (<i>Open Source</i>);
(I.4) Publicações correlacionadas à ciência da computação ou aplicadas na área de educação;
(I.5) Publicações dos últimos 5 anos.
Critérios de Exclusão (E)
(E.1) Publicações que fujam do tema abordado e que possuam mais de 5 anos de publicação;
(E.2) Trabalhos que não sejam artigos científicos, dissertações, teses ou qualquer tipo de trabalho acadêmico;
(E.3) Documentos duplicados;
(E.4) Publicações que tiveram menos de 10 citações.
Critérios de Qualidade (Q)
(Q.1) Definição clara dos objetivos da pesquisa;
(Q.2) Contexto adequado onde foi realizada a pesquisa;
(Q.3) Estratégia de extração de dados adequada aos objetivos da pesquisa;
(Q.4) Descrição dos resultados encontrados;
(Q.5) Evidência da contribuição para o aluno ou professor.

Fonte: Adaptado de Kitchenham, 2012.

3.5 RESULTADOS

Este trabalho constitui uma pesquisa bibliográfica sobre mineração de dados educacionais, empregando uma abordagem de revisão sistemática da literatura para uma análise do estado da arte. No total, foram utilizados 5 artigos para a realização das avaliações necessárias. A Figura 3.1 descreve as etapas do processo de pesquisa e seleção dos artigos de forma detalhada.

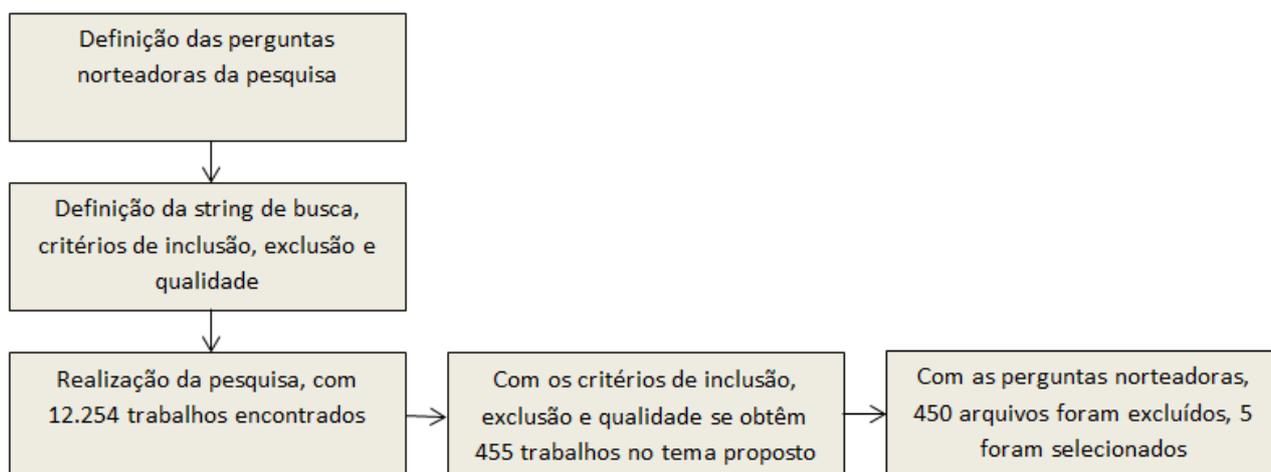


Figura 3.1: Etapas do processo de pesquisa e seleção de artigos

Fonte: Adaptado de Kitchenham, 2012.

Pode-se observar que na Tabela 3.5 é apresentado o total de artigos encontrados através dos critérios de inclusão, exclusão e qualidade, por cada site, tendo sido encontrados 6.134 no Periódicos CAPES e 6.120 no Scholar Google.

Tabela 3.5: Resultado das pesquisas realizadas

Base	Fase 1	Fase 2	Fase 3
Periódicos CAPES	6.134	234	1
Scholar Google	6.120	221	4
Total	12.254	455	5

Fonte: Dados da pesquisa.

O objetivo principal dos trabalhos de Taeza-Cruz e Capili-Kummer (2023), Guleria e Sood (2023), Cruz e Encarnacion (2021) e Repaso e Capariño (2020) foi analisar o desempenho dos alunos por meio da aplicação de técnicas de mineração de dados, seja utilizando ferramentas sintáticas ou desenvolvendo software com a mesma finalidade. Já o trabalho de Masci et al. (2018), o objetivo principal foi descobrir quais características dos estudantes estão associadas às pontuações dos testes do PISA. A Tabela 3.6 apresenta os artigos selecionados.

Tabela 3.6: Artigos selecionados

Autores	Ano de Publicação	Tema
TAEZA-CRUZ, Maria Elisa Linda; CAPILI-KUMMER, Marifel Grace	2023	Decision Support System to Enhance Students' Employability using Data Mining Techniques for Higher Education Institutions
GULERIA, Pratiyush; SOOD, Manu	2023	Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling
CRUZ, Maria Elisa Linda Taea; ENCARNACION, Riah Elcullada	2021	Analysis and Prediction of Students' Academic Performance and Employability Using Data Mining Techniques: A Research Travelogue
REPASO, Jennifer Anne A.; CAPARIÑO, Elenita T.	2020	Analyzing and Predicting Career Specialization using Classification Techniques
MASCI et al.	2018	Student and school performance across countries: A machine learning approach

Os artigos escolhidos passaram por uma leitura minuciosa e, em seguida, os dados relevantes foram extraídos. A discussão dos resultados foi conduzida por meio da organização de perguntas direcionadas à pesquisa. Logo após, os dados foram sintetizados e analisados para responder às questões de pesquisa. Nas seções 3.5.1 a 3.5.5 apresentam-se os resumos de cada

trabalho, na seção 3.5.6, apresentam-se as questões de pesquisa com as suas respectivas respostas e, na seção 3.5.7, apresentam-se as considerações finais do capítulo.

3.5.1 Decision Support System to Enhance Students' Employability using Data Mining Techniques for Higher Education Institutions (TAEZA-CRUZ; CAPILI-KUMMER, 2023)

Neste estudo, o objetivo foi estabelecer um sistema de suporte à decisão para instituições de ensino superior a fim de prever a empregabilidade dos estudantes. As técnicas de mineração de dados utilizadas puderam avaliar a preparação dos alunos para o emprego antes de concluírem seus estudos.

O estudo empregou o Scrum como metodologia de desenvolvimento e a linguagem de programação *Hypertext Preprocessor* (PHP) para criar o site. O software Weka também foi utilizado para criar um modelo de predição da empregabilidade dos estudantes. Especialistas em tecnologia da informação avaliaram o sistema desenvolvido por meio de um questionário online que avaliou a qualidade do sistema de acordo com os padrões ISO/IEC 25010.

O estudo revelou que o sistema de suporte à decisão desenvolvido obteve uma média geral de 4,43, classificada como "em grande medida", e está em conformidade com os Padrões de Qualidade de Software ISO 25010.

3.5.2 Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling (GULERIA; SOOD, 2023)

Neste estudo, o conceito de Aprendizado de Máquina é apresentado mostrando como ela aprende com experiências, inferências e concebe consultas complexas. Apresentaram ainda como técnicas de aprendizado de máquina podem ser usadas para desenvolver um *framework* educacional que recebe uma entrada de dados de estudantes, pais e, com inteligência, gera o resultado.

O *framework* integra as características de Aprendizado de Máquina (ML) e Inteligência Artificial Explicável (XAI) para analisar os fatores educacionais que são úteis para os estudantes conseguirem oportunidades profissionais e os ajuda a tomar a decisão certa para o crescimento em suas carreiras. Ele funciona como um sistema especialista com suporte de decisão para resolver

problemas da mesma forma que os humanos resolvem, ou seja, compreendendo, analisando e lembrando.

Os autores propuseram este *framework* para aconselhamento de carreira de estudantes usando técnicas de Aprendizado de Máquina e Inteligência Artificial. Modelos White Box e Black Box baseados em ML analisam o conjunto de dados educacionais, que inclui atributos acadêmicos e de empregabilidade importantes para oportunidades profissionais e capacitação dos estudantes.

No *framework* proposto, modelos White Box e Black Box foram treinados usando um conjunto de dados educacionais coletados no estudo. A pontuação das métricas obtidas pelo modelo Naive Bayes mereceram destaque.

3.5.3 Analysis and Prediction of Students' Academic Performance and Employability Using Data Mining Techniques: A Research Travelogue (CRUZ; ENCARNACION, 2021)

Neste estudo, foi realizada uma revisão abrangente e sistemática da literatura para identificar técnicas de mineração de dados, algoritmos e vários conjuntos de dados que levarão à predição inteligente da empregabilidade dos alunos. O mesmo método foi usado para determinar a relação entre o desempenho acadêmico e a empregabilidade dos alunos.

De acordo com as descobertas da pesquisa, as técnicas de mineração de dados mais frequentemente usadas para determinar o desempenho acadêmico dos alunos e a empregabilidade são as técnicas de Classificação, especificamente o algoritmo J48 (C4.5), o algoritmo Naïve Bayes e o algoritmo de Árvore de Decisão CHAID.

Os conjuntos de dados ou atributos mais frequentemente usados para prever o desempenho acadêmico dos alunos e a empregabilidade incluem a média acumulada de notas (CGPA - *Cumulative Grade Point Average*), gênero, habilidades técnicas, habilidades de comunicação, habilidades de resolução de problemas, habilidades analíticas, pensamento crítico e habilidades de tomada de decisão, atividades extracurriculares e idade, bem como fatores psicomotores, como comportamento, frequência e colocação em treinamento/estágio.

O desempenho acadêmico é o principal determinante da empregabilidade. A aplicação de técnicas de mineração de dados no meio acadêmico demonstrou seu valor na melhoria do

desempenho das instituições de ensino superior (IES). Como resultado, mais pesquisas são urgentemente necessárias para determinar a eficácia das abordagens, algoritmos e conjuntos de dados identificados como preditores da empregabilidade dos alunos. Além disso, abordagens automatizadas devem ser utilizadas para determinar sua precisão.

3.5.4 Analyzing and Predicting Career Specialization using Classification Techniques (REPASO; CAPARIÑO, 2020)

Neste estudo, foram exploradas várias técnicas de classificação na predição da especialização de graduados. Os conjuntos de dados usados foram obtidos dos graduados em Tecnologia da Informação do Campus Sarmiento da Bulacan State University de 2013 a 2016.

A partir desses dados, foi criado um modelo utilizando os algoritmos de classificação Naïve Bayes, J48, Random Forest e Support Vector Machine, com 18 atributos. Entre os modelos criados, os algoritmos Naïve Bayes e Random Forest apresentaram uma melhor classificação de precisão e valores aceitáveis de ROC (Curva Característica de Operação do Receptor) e RMSE (Erro Quadrático Médio).

O desempenho dos estudantes por área de disciplina também foi determinado, e com base nisso, os estudantes tiveram desempenho satisfatório tanto em habilidades técnicas quanto em habilidades interpessoais, como evidenciado pelo desempenho altamente satisfatório na disciplina de Estágio.

No entanto, houve poucos graduados que seguiram carreiras na área de Redes de Computadores, e medidas devem ser tomadas para melhorar o desempenho nas disciplinas iniciais de Programação e Redes.

3.5.5 Student and school performance across countries: A machine learning approach (MASCI et al., 2018)

Neste estudo, foram desenvolvidos e aplicados novos métodos estatísticos e de aprendizado de máquina para analisar os determinantes das pontuações dos testes do PISA 2015 de estudantes em nove países: Austrália, Canadá, França, Alemanha, Itália, Japão, Espanha, Reino Unido e Estados Unidos.

O objetivo foi descobrir quais características dos estudantes estão associadas às pontuações dos testes do PISA e quais características das escolas estão associadas ao valor agregado da escola (medido no nível da escola). Um objetivo específico dessa abordagem é explorar não linearidades nas associações entre covariáveis e pontuações dos testes, bem como modelar interações entre fatores do nível da escola que afetam os resultados.

Para abordar essas questões, foi aplicada uma metodologia de dois estágios usando métodos flexíveis baseados em árvores. Primeiro, executaram árvores de regressão multinível no primeiro estágio para estimar o valor agregado da escola. No segundo estágio, relacionaram o valor agregado da escola estimado com variáveis do nível da escola por meio de árvores de regressão e modelos de impulso.

Os resultados mostraram que, embora várias características de estudantes e escolas estejam significativamente associadas ao desempenho dos estudantes, existem diferenças marcantes entre os países. A abordagem proposta permite uma descrição aprimorada das funções de produção educacional estruturalmente diferentes em diferentes países.

3.5.6 Respostas das questões de pesquisa

Q1: É passível de automação um sistema de avaliação e predição de desempenho de estudantes deterministicamente com previsibilidade de desempenho profissional?

Taeza-Cruz e Capili-Kummer (2023) desenvolveram um sistema para o ensino superior com foco na empregabilidade dos estudantes e na oferta de uma educação de qualidade que produza graduados qualificados para oportunidades de emprego. Este sistema foi considerado útil e necessário, dado o compromisso da Universidade de Nizwa com a empregabilidade dos estudantes. O uso do sistema desenvolvido simplificaria a identificação de estudantes em risco por parte dos orientadores acadêmicos e permitiria a eles fornecer intervenções rápidas e apropriadas para ajudar seus orientados a melhorar seu desempenho acadêmico. Este estudo mostrou que modelos determinísticos, como o modelo Árvore de Decisão J48, é uma das técnicas mais apropriadas de mineração de dados para prever a empregabilidade de estudantes, onde este obteve uma precisão de 96,6135%. Os resultados sugerem que a situação acadêmica, a nota do estágio e a carga horária são as variáveis mais significativas da empregabilidade dos alunos, conforme a Tabela 3.7.

Tabela 3.7: Variáveis significativas da empregabilidade

Variável	Descrição	Valores
Situação Acadêmica	Classificação do CGPA para o grau de bacharel.	Excelente com Honras, Excelente, Muito Bom, Bom, Satisfatório e Em Período de Avaliação.
Nota do Estágio	A nota final recebida pelo estudante durante o estágio.	A (4.0); A- (3.7); B+ (3.3); B (3.0); B- (2.7); C+ (2.3); C (2.0); C- (1.7); D+ (1.3); D (1.0); F (0).
Carga horária	Horas de crédito completadas pelo estudante.	9 – 160 horas.

Fonte: Adaptada de TAEZA-CRUZ; CAPILI-KUMMER, 2023.

Guleria e Sood (2023) desenvolveram um *framework* onde este utiliza técnicas de aprendizado de máquina para realizar previsões de fatores acadêmicos e de oportunidades de emprego como parte do modelo de orientação profissional proposto. Os modelos de aprendizado de máquina precisam ser treinados com os conjuntos de dados educacionais coletados em várias instituições educacionais, com um maior número de atributos para avaliação dos estudantes, levando em consideração fatores como oportunidades de emprego com o objetivo de aprimorar a empregabilidade e o desenvolvimento de habilidades. No entanto, o escore de Recall e F-Measure alcançado pelo Naive Bayes para realizar previsões é de 91,2% e 90,7%, que é o melhor em comparação com os escores dos modelos de Regressão Logística, Árvore de Decisão, SVM, KNN e Ensemble considerados no estudo. Algumas variáveis preditoras foram utilizadas, como notas do ensino médio, notas do ensino superior, experiência de trabalho, nota em teste eletrônico de empregabilidade e notas de curso de especialização (se houver), conforme Tabela 3.8.

Tabela 3.8: Variáveis preditoras

Variável	Tipo de Dados	Exemplo
Notas (ensino médio, ensino superior, teste de empregabilidade, especialização (se houver))	Float	78.33, 73.6, 59.43
Experiência de Trabalho	Boolean	Yes/No

Fonte: Adaptada de GULERIA; SOOD, 2023.

Cruz e Encarnacion (2021) realizaram uma revisão da literatura, onde vários estudos mostraram resultados notáveis na predição da empregabilidade dos estudantes. As técnicas de classificação foram predominantemente usadas para tarefas de predição e análise, especificamente o uso dos algoritmos J48 (C4.5), Naïve Bayes e algoritmos de Árvore de Decisão CHAID, devido aos seus altos resultados de precisão. O estudo revelou algumas variáveis comuns que foram usadas para classificação, como média acumulada de notas (CGPA), gênero, idade, grupo de estudo e aulas particulares, conforme Tabela 3.9. Além disso, a revisão sistemática da literatura e estudos relacionados demonstrou que o desempenho acadêmico influencia significativamente a empregabilidade dos estudantes.

Tabela 3.9: Variáveis relacionadas ao estudante

Variável	Descrição	Valores
Gênero	Gênero do estudante	Masculino, Feminino.
Comunidade	Status social, conforme classes sociais da Índia.	SC (Castas Programadas), ST (Tribos Programadas), MBC (Classes Mais Atrasadas), BC (Classes Atrasadas), OC (Outras).
Educação dos pais	Indica se os pais foram alfabetizados.	Ambos educados, mãe educada, pai educado, ambos não educados.
Localidade de Moradia	Localidade onde mora o estudante.	Empresarial, Município, Zona Rural.
Transporte	Modo de transporte para escola.	A pé, bicicleta, ônibus coletivo, ônibus escolar, automóvel próprio.
Tipo de Escola	Determina o tipo de escola que o aluno estudou o nível superior.	Coeducação, Meninos, Meninas.
Aulas Particulares	A maioria dos pais envia seus filhos para aulas particulares.	Sim, Não.
Grupo de Estudo	Seis tipos de grupos de estudo, baseados em disciplinas básicas, são oferecidos no nível superior.	Primeiro, segundo, terceiro, quarto, quinto, sexto.
Localidade da Escola	Localidade da escola do estudante.	Empresarial, Município, Zona Rural.
Acesso à celular	Indica se o estudante possui celular.	Sim, Não.
Acesso a esportes	Indica se o aluno possui acesso a esportes e lazer.	Sim, Não.

Acesso a computador	Indica se o aluno possui acesso a computador.	Sim, Não.
Acesso à internet	Indica se o aluno possui acesso à internet.	Sim, Não.
Profissão do pai	Indica a ocupação do pai do estudante.	Trabalhador braçal, Agricultor, Tecelão, Emprego no setor privado, Emprego no setor público, Empresário, Não aplicável.
Profissão da mãe	Indica a ocupação da mãe do estudante.	Dona de Casa, Trabalhadora braçal, Agricultora, Tecelã, Emprego no setor privado, Emprego no setor público, Não aplicável.
Salário dos pais	Indica a soma dos salários dos pais.	0..0.9K, 1K..2.9K, 3K...4.9K, 5K..9K, 10K..20K, acima de 20K, Não aplicável.
Média acumulada de notas (CGPA)	Notas obtidas no nível superior. É dividido em sete classes.	O – 90% a 100%, A – 80% - 89%, B – 70% - 79%, C – 60% - 69%, D – 50% - 59%, E – 40% - 49%, F - < 40%.

Fonte: Adaptada de CRUZ; ENCARNACION, 2021.

Repaso e Capariño (2020) revelaram que algoritmos de classificação, como Naïve Bayes, podem ser apropriados para prever a especialização de carreira de um graduado. Cada disciplina cursada habilita o aluno a uma competência técnica e também a uma habilidade interpessoal. O estudo apresenta algumas habilidades interpessoais como: Resolução de problemas, criatividade, trabalho em equipe, utilização estratégica de tecnologia da informação e planejamento, conforme Tabela 3.10. Os estudantes do primeiro ano demonstraram um desempenho razoável em sua primeira disciplina de Programação e Redes, tiveram resultado satisfatório em disciplinas, como Análise e Projeto de Sistemas, Projeto de Conclusão de Curso e Estágio. Nesse sentido, o departamento de tecnologia da informação do campus e os professores puderam implementar programas para melhorar o desempenho dos estudantes, especialmente no primeiro ano, especificamente em Programação e em algumas disciplinas do segundo e terceiro ano, como Redes e Engenharia de Software. Além disso, as boas práticas do departamento foram mantidas para sustentar o desempenho dos estudantes, especialmente em disciplinas de Comunicação e Ética.

Tabela 3.10: Atributos e habilidades correspondentes

Disciplina/Atributo	Habilidades Interpessoais	Valores
Programação	Resolução de problemas, criatividade.	VG (1.0, 1.25, 1.5), G (1.75, 2.0, 2.25), F (2.5, 2.75, 3.0)
Banco de Dados	Resolução de problemas, criatividade.	
Desenvolvimento Web	Resolução de problemas, criatividade.	
Redes de Computadores	Resolução de problemas, criatividade.	
Análise e Projeto de Sistemas	Resolução de problemas, criatividade, trabalho em equipe, utilização estratégica da tecnologia da informação e planejamento.	
Engenharia de Software	Resolução de problemas, criatividade.	
Sistemas Operacionais	Resolução de problemas, criatividade.	
Comunicação	Habilidade de comunicação.	
Ética Profissional	Trabalho em equipe, ética de trabalho.	
Gerenciamento de Projetos	Resolução de problemas, criatividade, trabalho em equipe.	
Estágio	Trabalho em equipe.	

Fonte: Adaptada de REPASO; CAPARIÑO, 2020.

Q2: Quais tecnologias atuais poderiam ser usadas em um protótipo nesse sentido?

Taeza-Cruz e Capili-Kummer (2023) empregaram métodos descritivos e de desenvolvimento, incluindo métodos ágeis Scrum e a linguagem de programação *Hypertext Preprocessor* (PHP) para criar o site.

Guleria e Sood (2023) realizaram a experimentação usando a caixa de ferramentas de aprendizado de máquina do MATLAB, como por exemplo, na avaliação de desempenho dos modelos de aprendizado de máquina caixa branca e caixa preta.

Nos estudos de Taeza-Cruz e Capili-Kummer (2023), Guleria e Sood (2023) e Repaso e Capariño (2020) foi predominante o uso do *framework* Weka. Cruz e Encarnacion (2021) em sua revisão da literatura também encontraram trabalhos que utilizaram esse software para criar modelos de predição. A Tabela 3.11 apresenta os artigos selecionados e um resumo das respostas de forma sintética para as questões de pesquisa 1 e 2:

Tabela 3.11: Resultados dos artigos selecionados

Estudo	Decision Support System to Enhance Students' Employability using Data Mining Techniques for Higher Education Institutions (TAEZA-CRUZ; CAPILI-KUMMER, 2023)	Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling (GULERIA; SOOD, 2023)	Analysis and Prediction of Students' Academic Performance and Employability Using Data Mining Techniques: A Research Travelogue (CRUZ; ENCARNACION, 2021)	Analyzing and Predicting Career Specialization using Classification Techniques (REPASO; CAPARIÑO, 2020)
Sistema	Predição de empregabilidade dos estudantes	Aconselhamento de carreira de estudantes	Revisão da literatura	Predição de especialização de graduados
Nível de escolaridade	graduação	graduação	graduação	graduação
Origem	Omã	Tailândia	Diversas	Filipinas
Framework	Weka	Weka	Diversos	Weka
Técnica de melhor acurácia	Árvore de Decisão J48	Naive Bayes	Árvore de Decisão J48	Naïve Bayes e Random Forest
Percentual	96,6135%	89,2%	-	65%

Fonte: Dados da pesquisa

Q3: Ranks (como TOEFL, GRE, ENEM, PISA) são eficazes na elaboração de diretrizes e guias para as instituições de ensino terem melhores egressos?

Masci et al. (2018) aplicaram novos métodos de análise, retirados da literatura de aprendizado de máquina, para examinar os determinantes das notas dos alunos no PISA e do valor agregado das escolas. Os resultados confirmaram muitos dos relacionamentos que já eram até então conhecidos a partir da análise estatística, mas fornecem uma compreensão nova e enriquecida de como as não-linearidades entre os fatores quanto as interações entre eles determinam o desempenho educacional. Essas percepções derivam do reconhecimento de que o processo educacional é complexo, desconhecido em seus mecanismos específicos e heterogêneo entre os países participantes da pesquisa. Uma parte do estudo mostra que as variáveis relacionadas aos alunos conseguem explicar parte da variabilidade de seu desempenho: Índice socioeconômico, ansiedade, motivação, gênero e educação dos pais são algumas das variáveis mais influentes. Sua associação com as notas do PISA e sua capacidade de explicar a variabilidade no sucesso acadêmico dos alunos diferem substancialmente entre os países pesquisados. A porcentagem da variabilidade do sucesso acadêmico dos alunos explicada no valor agregado das escolas também varia entre os países. Os países nos quais a variância estimada do valor agregado das escolas é alta são caracterizados pela heterogeneidade no nível da escola. Pelo contrário, os países nos quais a variância do valor agregado das escolas é limitada em magnitude oferecem uma experiência mais homogênea entre as escolas. Existem claras implicações de política ao notar, por exemplo, que a relação entre alunos e professores tem uma influência relativamente alta no Canadá, Japão e Espanha, mas não em outros lugares. Em muitos países, as ações que podem melhorar de forma mais eficaz os resultados educacionais não são políticas educacionais em si, mas sim políticas sociais.

3.5.7 Considerações Finais do Capítulo

Diante das informações obtidas neste capítulo, podem ser definidas as variáveis necessárias para a realização das predições pelo aplicativo PredictApp, a seguir:

- Obtidas a partir do sistema acadêmico: Notas, faltas, coeficiente de rendimento.
- Obtidas a partir do preenchimento de questionário pelo estudante: Experiência de Trabalho, Educação dos pais, Localidade de Moradia, Transporte, Tipo de Escola, Aulas Particulares, Grupo de Estudo, Escola, Acesso à internet, Acesso a esporte/lazer e Renda Familiar.

4 DESENHO DA ARQUITETURA

Os modelos de classificação só terão utilidade quando os estudantes interagirem com eles, uma vez que os modelos visam prever o resultado de seu desempenho, o que é crucial do ponto de vista de suas carreiras. Diferentes modelos de classificação podem ser utilizados para classificar dados de estudantes na busca de obter os melhores resultados. A Figura 4.1 apresenta o modelo conceitual proposto para o aplicativo PredictApp.

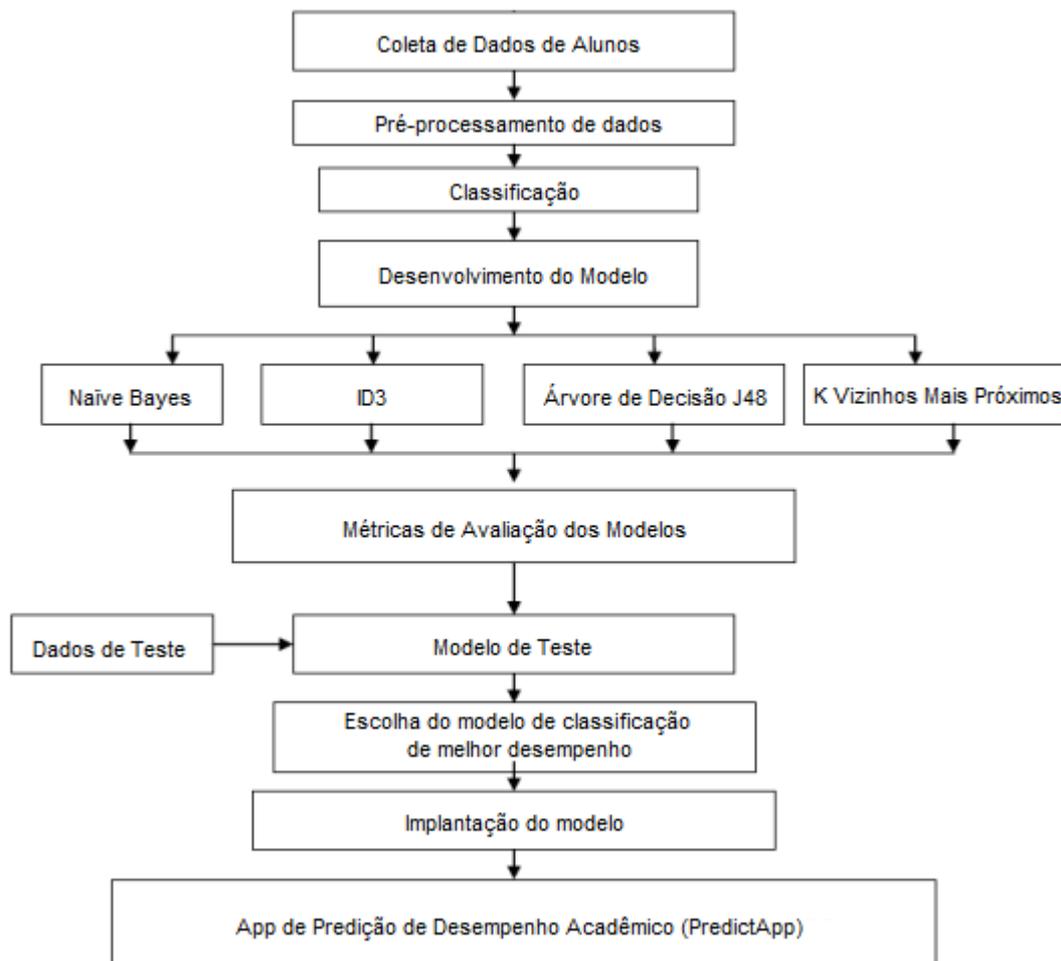


Figura 4.1: Modelo Conceitual PredictApp

O modelo será implementado utilizando a linguagem de programação Python. O aplicativo para dispositivos móveis será desenvolvido e irá consumir uma API REST para fornecer aos usuários as predições. Para esse fim, é apresentada a arquitetura para a implantação do aplicativo, conforme a Figura 4.2.

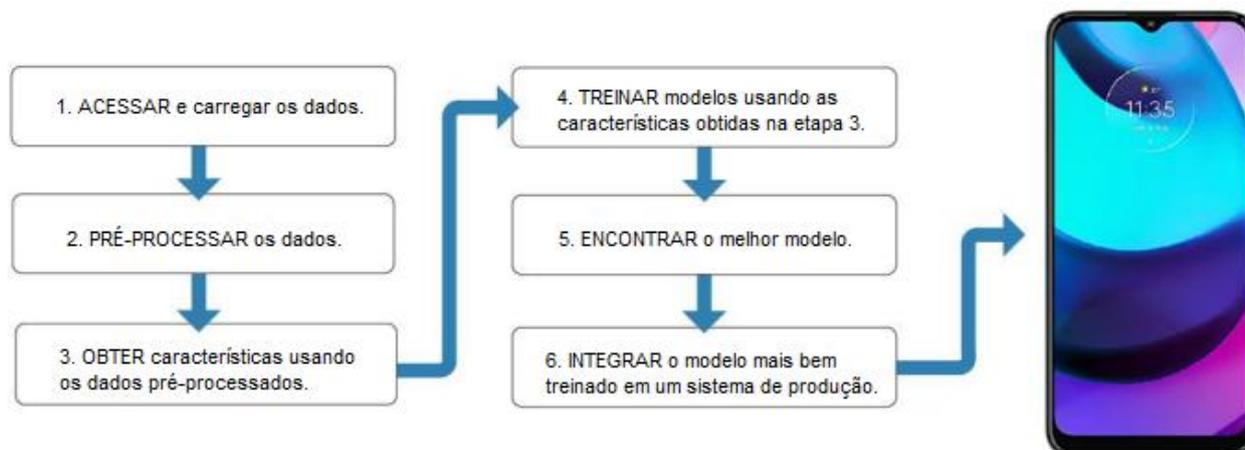


Figura 4.2: Processo de implantação do modelo

4.1 ETNOGRAFIA

Etnografia é uma técnica de observação que pode ser usada para compreender os processos operacionais e ajudar a extrair os requisitos de apoio para esses processos. Um analista faz uma imersão no ambiente de trabalho em que o sistema será usado. O trabalho do dia a dia é observado e são feitas anotações sobre as tarefas reais em que os participantes estão envolvidos. O valor da etnografia é que ela ajuda a descobrir requisitos implícitos do sistema que refletem as formas reais com que as pessoas trabalham, em vez de refletir processos formais definidos pela organização [Sommerville, 2019].

4.1.1 Análise das Tarefas

As atividades (tarefas) do usuário deverão ser compatíveis com o sistema de computador. Sem uma compreensão do que o usuário deseja fazer com o sistema de computador não há nenhuma perspectiva realista de projetar uma interface com o usuário que seja eficaz [Sommerville, 2019].

A análise da tarefa identifica a forma que o usuário utiliza para alcançar o objetivo final e as possíveis circunstâncias encontradas. Assim, entende-se como o usuário realiza suas tarefas no mundo real.

Na Figura 4.3 são mostradas as tarefas executadas pelo professor no procedimento de inserir notas no portal acadêmico:

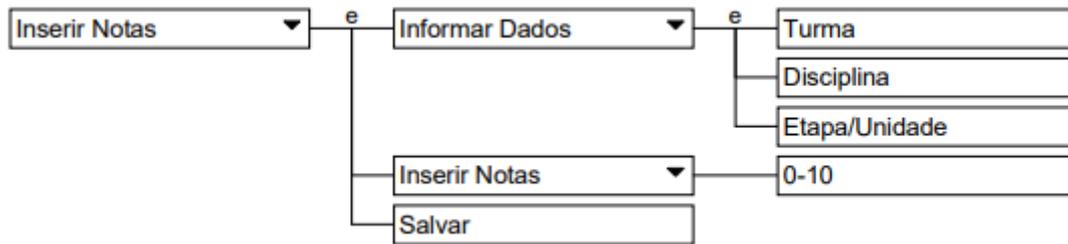


Figura 4.3: Tarefa Inserir Notas
Fonte: Dados da pesquisa

Agente: Professor.

Objetivo: Inserir notas no portal acadêmico.

Tarefa: O professor deve informar os dados (turma, disciplina, etapa/unidade) para que o sistema possa carregar a lista com os nomes dos alunos. Em seguida, o professor deve inserir as notas de cada aluno, atribuindo notas entre 0 a 10. Logo após, o professor deve clicar no botão Salvar.

Ação (último passo): Salvar

Objeto: Dados.

Ambiente: Os dados são gravados no banco de dados. A confirmação se dará quando aparecer uma mensagem.

Na Figura 4.4 são mostradas as tarefas executadas pelo professor no procedimento de inserir faltas no portal acadêmico:



Figura 4.4: Tarefa Inserir Faltas
Fonte: Dados da pesquisa

Agente: Professor.

Objetivo: Inserir faltas no portal acadêmico.

Tarefa: O professor deve informar os dados (turma, disciplina, etapa/unidade, data) para que o sistema possa carregar a lista com os nomes dos alunos. Em seguida, o professor deve marcar o *checkbox* para o aluno que esteve ausente naquela data. Logo após, o professor deve clicar no botão Salvar.

Ação (último passo): Salvar.

Objeto: Dados.

Ambiente: Os dados são gravados no banco de dados. A confirmação se dará quando aparecer uma mensagem.

Na Figura 4.5 são mostradas as tarefas executadas pelo aluno no procedimento de consultar notas e faltas no portal acadêmico:

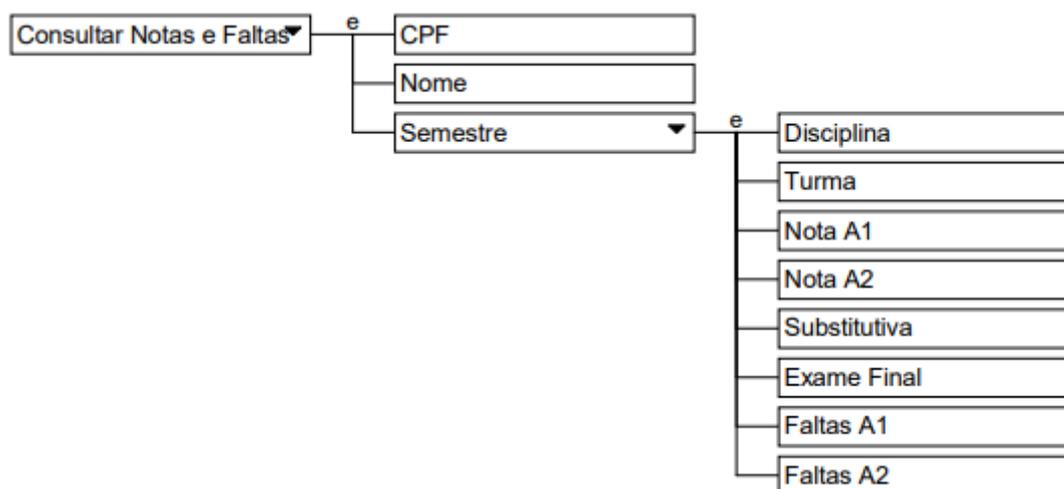


Figura 4.5: Tarefa Consultar Notas e Faltas
Fonte: Dados da pesquisa

Agente: Aluno.

Objetivo: Consultar notas e faltas no portal acadêmico.

Tarefa: O aluno deve informar o semestre para que o sistema possa carregar os dados das disciplinas como: Turma, Nota a1, Nota a2, Substitutiva, Exame Final, Faltas A1 e Faltas A2. Cada vez que o aluno selecionar um semestre diferente, o sistema deverá carregar os dados relacionados às disciplinas cursadas no semestre selecionado.

Ação (último passo): Selecionar semestre.

Objeto: Dados.

Ambiente: Os dados são retornados do banco de dados e exibidos na tela.

Após o estudo dessa técnica, é possível criar representações, mesmo que superficiais, para a interface com o usuário. Fica mais fácil criar uma interface onde o usuário identifique a sequência dos passos na realização da tarefa.

As técnicas utilizadas neste capítulo não são totalmente suficientes para uma análise completa dos requisitos do sistema de integração das aplicações. Existem outras técnicas bastante aceitas como a FAST (*Facilitated Application Specification Techniques*) que podem ser utilizadas

em conjunto. As técnicas mostradas servem como exemplo de técnicas de levantamento de requisitos, expondo sua utilidade no processo de descoberta e análise dos mesmos no assunto em estudo.

4.2 ESPECIFICAÇÃO DE REQUISITOS

A função e o desempenho atribuídos ao software como parte da engenharia do sistema são refinados ao se estabelecer uma descrição completa da informação, uma indicação dos requisitos de desempenho e restrições de projeto, critérios de validação apropriados e outros dados pertinentes aos requisitos [Pressman, 2016].

Nesta seção, trata-se de estabelecer as funções e restrições requeridas pelo sistema. Sabe-se que é uma atividade particularmente importante, que os erros produzidos nessa fase inevitavelmente produzem problemas posteriores no projeto e na implementação do sistema.

Os requisitos apresentados a seguir foram descritos através do uso das técnicas de levantamento de requisitos, como entrevistas e observações.. No Apêndice A, segue formulário de perguntas utilizado nas entrevistas com os usuários envolvidos.

4.2.1 Requisitos Funcionais

- O aplicativo não realizará cadastro de usuário, somente consultas de informações. Os cadastros devem ser realizados através do portal acadêmico.
- Deverão ser exibidos os dados CPF e nome do aluno na tela de exibição do boletim.
- O aluno deverá visualizar apenas as disciplinas do semestre selecionado com suas respectivas informações e ao logar o app carregará a tela com os dados do semestre atual.
- O aluno não irá digitar o semestre no qual deseja visualizar as notas. O aplicativo deverá trazer essa informação do banco de dados para que ele apenas a selecione.
- Quando selecionado um semestre que o aluno não tenha cursado nenhuma disciplina, o aplicativo deverá exibir uma mensagem informando o ocorrido.

- O aluno deverá preencher um questionário com perguntas no aplicativo para que o modelo (algoritmo) de MDE possa realizar a predição de desempenho acadêmico e retornar as informações para o aplicativo.

4.2.2 Requisitos Não Funcionais

- Eficiência: o tempo de recepção das informações deve ser realizado da forma mais rápida possível;
- Confiabilidade: garantir que os dados apresentados no aplicativo são realmente do respectivo aluno;
- Usabilidade: a interface deverá ser de fácil compreensão para o aluno;
- Portabilidade: a integração com a base de dados será independente de hardware/software;
- Manutenibilidade: caso haja alterações, como inclusão de mais dados a serem informados pela instituição de ensino, o aplicativo deverá ser atualizado o mais breve possível;
- Transportabilidade: a integração com a base de dados deve ser desenvolvida de tal forma que o aluno possa acessar o aplicativo e consultar suas notas/faltas e predição de desempenho de qualquer região do país;
- Compreensibilidade: a forma de utilização do aplicativo deve ser explicada para o aluno em forma de tutorial em texto ou vídeo;
- Implantabilidade: a integração do aplicativo com o portal acadêmico e com o modelo de predição deve ser implantada, observando o compartilhamento de dados entre os sistemas, possibilitando uma comparação de resultados;
- Integridade: ao abrir o aplicativo, deverá ser solicitada uma forma de autenticação do usuário para que o mesmo tenha acesso.

A análise de requisitos é o primeiro passo técnico do processo de engenharia de software. É nesse ponto que uma declaração geral do escopo do software é aprimorada numa especificação

concreta que se torna a base para todas as atividades de engenharia de software que se seguirão [Pressman, 2016].

Não há um padrão a ser seguido quanto à ordem das atividades no processo de requisitos, mesmo porque a análise é contínua, novos requisitos vão surgindo ao longo do processo e mesmo após a entrega do produto de software.

Ao final desta seção, conclui-se a atividade de especificação de requisitos, onde já se têm informações fundamentais para o desenvolvimento do aplicativo. Apresentam-se nas próximas seções detalhes técnicos da arquitetura do aplicativo com diversos diagramas, cada um com as suas particularidades.

4.3 ARQUITETURA DA INTEGRAÇÃO

Geralmente, sistemas acadêmicos são compreendidos com facilidade entre os usuários, mas, todavia com as novidades que surgem em tecnologia da informação, pontos importantes podem ser melhorados. Um exemplo disso é a implantação de um aplicativo para que os alunos possam consultar suas notas, frequências e predição de desempenho acadêmico, de maneira que os problemas abordados neste trabalho sejam solucionados.

Instituições de ensino encontram dificuldade para disponibilizar uma forma para que seus alunos consigam obter informações acadêmicas a partir de um dispositivo móvel, visto que até mesmo quando o portal acadêmico dessas instituições é desenvolvido utilizando design responsivo, nem sempre a interface se ajusta e oferece usabilidade fluida.

A tecnologia de Web Services surgiu com o propósito de facilitar a integração de sistemas distintos, resolvendo problemas de comunicação entre softwares. Com base nesta situação tem-se como proposta desse trabalho o estudo das tecnologias existentes atualmente para o desenvolvimento de um aplicativo para dispositivos móveis que realize a integração com a base de dados do portal acadêmico, visando à aplicação de Web Services no aspecto da interoperabilidade. Esta integração ocorre através da possibilidade do aplicativo para dispositivos móveis utilizar as informações da base de dados do portal acadêmico, trazendo uma nova forma que irá facilitar o acesso do aluno na obtenção de informações acadêmicas e melhorando ainda mais o relacionamento da instituição de ensino com o mesmo.

As informações contidas no banco de dados devem ser apresentadas na tela do aplicativo sem nenhuma divergência. Para isso, é montada uma arquitetura onde o aplicativo faz uso de uma API (*Application Programming Interface*) de desenvolvimento que, assim que o usuário realizar a ação de consulta, a API realiza a solicitação do serviço do Web Service. Este por sua vez, realiza uma consulta SQL (*Structured Query Language*) ao banco de dados que retorna ao Web Service o resultado da consulta. Por último, o Web Service envia a resposta ao aplicativo com as informações solicitadas, de forma transparente para o usuário, como se o mesmo estivesse utilizando o próprio portal acadêmico.

Web Service é uma tecnologia que desempenha um papel importante na comunicação entre aplicativos, incluindo a comunicação entre *smartphones* e bancos de dados. Web Service é um serviço que permite que diferentes aplicativos se comuniquem e troquem dados pela internet, independentemente da plataforma ou linguagem de programação utilizada por cada aplicativo. Ele atua como intermediário para permitir que sistemas heterogêneos se comuniquem de maneira padronizada e interoperável [Ort, 2005].

A Figura 4.6 mostra como um dispositivo móvel acessa informações de um banco de dados, ou seja, consumindo serviços de um web service:

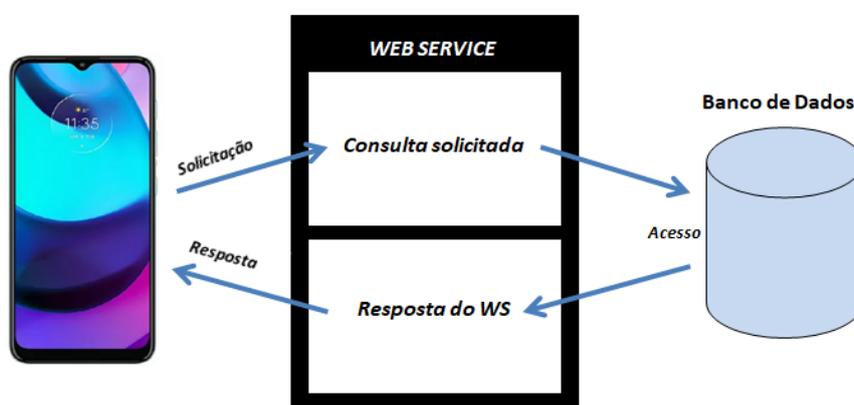


Figura 4.6: Smartphone consumindo serviço

Na comunicação entre um *smartphone* e um banco de dados, um web service pode ser utilizado de diversas formas, dentre elas [Lecheta, 2015]:

- Exposição de Funcionalidades do Banco de Dados: O banco de dados é geralmente hospedado em um servidor, e acessá-lo diretamente do *smartphone* pode ser arriscado em termos de segurança e eficiência. Um web service pode ser desenvolvido no servidor para

expor as funcionalidades específicas do banco de dados que o aplicativo no *smartphone* precisa acessar.

- Integração com Aplicativos Móveis: O aplicativo no *smartphone* faz chamadas ao web service usando as APIs definidas. O web service processa essas chamadas, interage com o banco de dados e retorna as respostas para o aplicativo no formato adequado.

4.4 DIAGRAMAS

Nesta seção apresentam-se os diagramas em UML (*Unified Modeling Language*), onde se descreve a organização do sistema de integração dos softwares. O sistema é representado através de um conjunto de diagramas, onde cada diagrama se refere a uma visão parcial do mesmo e o conjunto forma um todo integrado e coerente.

4.4.1 Diagrama de Casos de Uso

O diagrama de casos de uso descreve o sistema do ponto de vista do usuário, como o usuário interage com o sistema sob determinadas circunstâncias. Ele representa graficamente esta interação, e define o contexto do sistema. Neste trabalho, o professor pode executar algumas ações no sistema, como: Inserir Notas, Inserir Faltas e Consultar Predição. Consequentemente, o aluno executa outras ações, como: Consultar Notas, Consultar Faltas, Responder *Survey* e Consultar Predição. Apresenta-se na Figura 4.7 o diagrama de casos de uso simplificado, os quais demonstram as ações que o professor e o aluno podem executar no sistema.

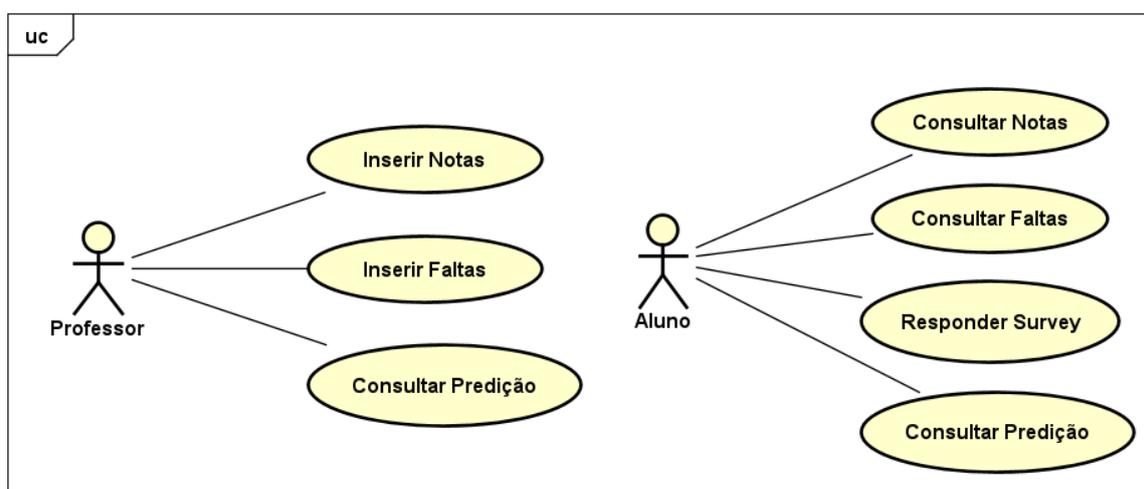


Figura 4.7: Casos de Uso

4.4.2 Diagrama de Atividades

A consulta das notas, faltas e previsão de desempenho acadêmico acontecem da seguinte forma: o professor informa os dados da turma, digitando as notas e faltas. O aluno por sua vez, consulta as notas/faltas, responde as perguntas e o sistema retorna as informações de previsão de desempenho acadêmico do aluno, conforme Figura 4.8.

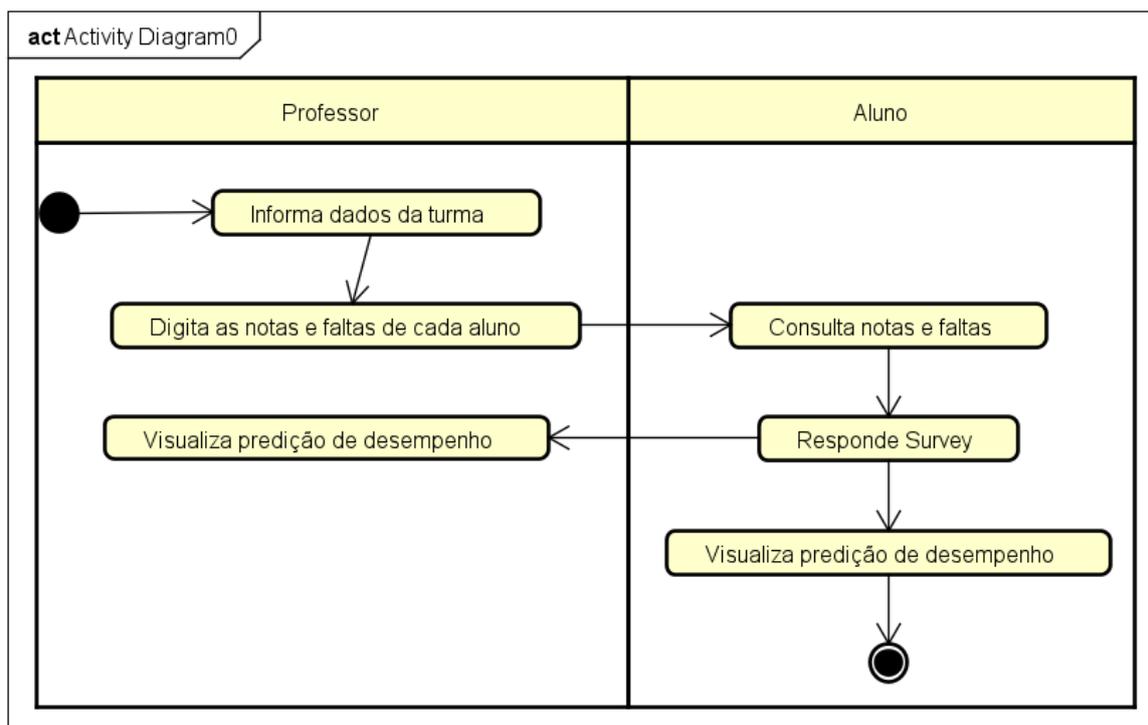


Figura 4.8: Diagrama de atividades

4.4.3 Diagrama de Pacotes

O diagrama de pacotes mostra as camadas/subsistemas do sistema de integração das aplicações. Cada subsistema representa uma coleção de objetos e outro(s) pacote(s), ou seja, cada subsistema inclui uma série de outros componentes. Esses componentes estão representados de forma muito abstrata, onde não incluem muitos detalhes das interações.

O contexto do sistema de integração das aplicações está dentro de um subsistema, que se ocupa de executar os serviços solicitados, além de outros subsistemas que juntos formam o sistema de integração das aplicações.

Na Figura 4.9 tem-se o diagrama de pacotes de toda a estrutura:

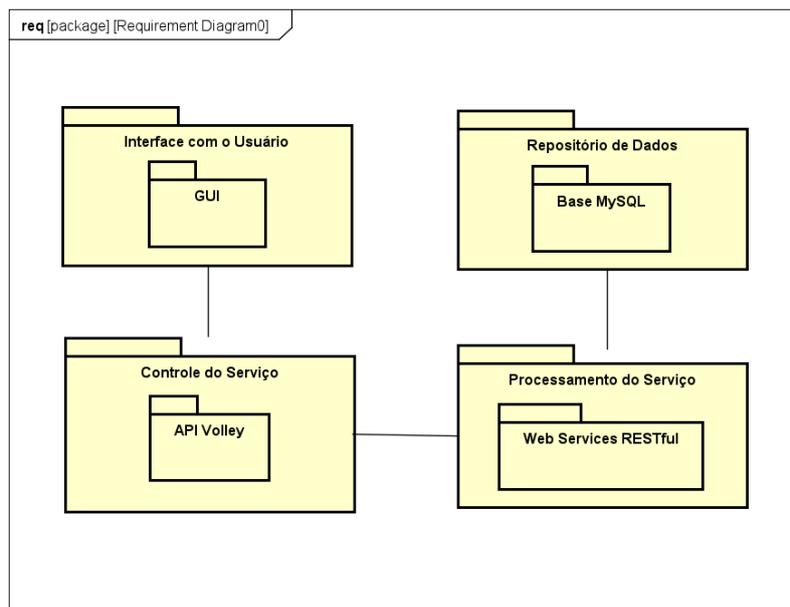


Figura 4.9: Diagrama de Pacotes

Conforme a figura 5.7, têm-se as camadas assim identificadas:

- Interface com o Usuário – os objetos se ocupam da apresentação de informações para o usuário.
- Controle de Serviço – os objetos se ocupam de controlar as informações.
- Processamento do Serviço – os objetos se ocupam de executar o serviço solicitado.
- Repositório de Dados – os objetos se ocupam do armazenamento de dados para futuro processamento.

4.4.4 Diagrama de Classes

Uma classe de objeto pode ser pensada como uma definição geral de um tipo de objeto do sistema. Uma associação é um link entre classes que indica algum relacionamento entre essas classes. Conseqüentemente, cada classe pode precisar de algum conhecimento sobre sua classe associada [Sommerville, 2019].

Tem-se então a GUI formada pelas classes LoginActivity e NotasActivity, indicando que essas classes interagem com as classes Web Services: getUsuarios, getSemestres e getNotas, para solicitar seus serviços e obter os resultados que elas precisam para cumprir os objetivos do aplicativo. Estas três classes Web Services ainda interagem com a classe Web Service dbConnection que possui os dados de conexão com o banco de dados. Obviamente as classes Web

Services ficam localizadas no servidor em nuvem do Heroku, que gerencia os acessos, a segurança e outros recursos.

O diagrama de classes da Figura 4.10 descreve as classes responsáveis por realizar a integração dos sistemas e suas relações.

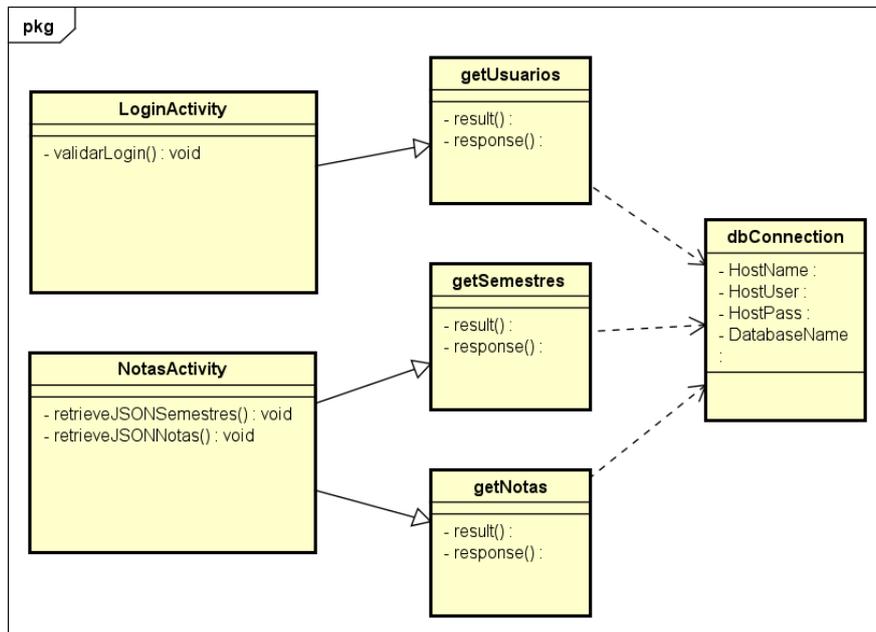


Figura 4.10: Diagrama de Classes

4.4.5 Diagramas de Sequência

Como o nome indica, um diagrama de sequência mostra a sequência de interações que ocorrem durante um caso de uso em particular ou em uma instância de caso de uso [Sommerville, 2019].

Na operação de Logar no Aplicativo, o aluno informa os dados de CPF e senha, a API *Volley*⁵ se encarrega de enviar estes dados para o Web Service `getUsuarios`, que realiza a consulta no banco de dados e retorna o resultado com o usuário logado com sucesso, conforme a Figura 4.11.

⁵ <https://developer.android.com/training/volley?hl=pt-br>

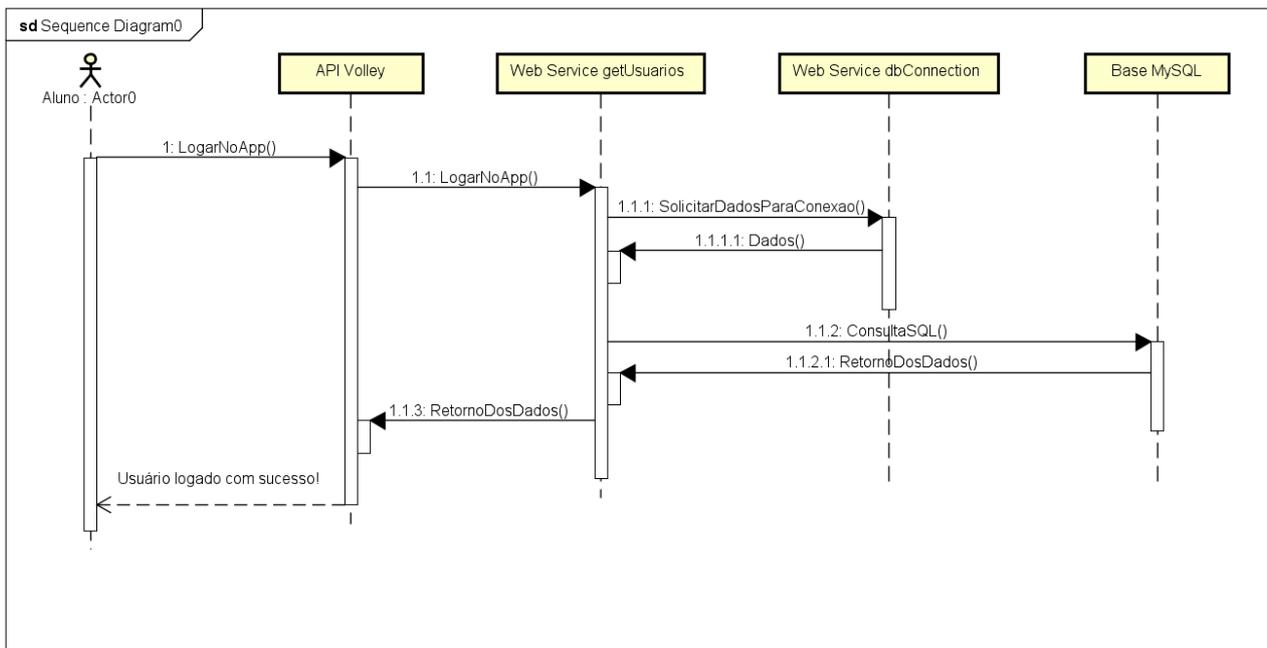


Figura 4.11: Diagrama de Sequência LogarNoApp

O diagrama de sequência mostra a colaboração dinâmica entre os vários objetos do sistema, apresentando a sequência de ações executadas por ele e a sequência de mensagens enviadas entre os objetos.

Os diagramas de sequência em UML são usados, principalmente, para modelar as interações entre os atores e os objetos em um sistema e as interações entre os próprios objetos. A UML tem uma sintaxe rica para diagramas de sequência, que permite a modelagem de vários tipos de interação [Sommerville, 2019].

Na operação de Carregar *Spinner*⁶ dos Semestres, o aplicativo possui um requisito de carregar um *Spinner* com os semestres cadastrados pela instituição de ensino. Nesse caso, esse Web Service entra em ação logo que o aluno consegue logar no aplicativo com sucesso. Então, a API *Volley* se encarrega de enviar estes dados para o Web Service *getSemestres*, que realiza a consulta no banco de dados e retorna o resultado, conforme a Figura 4.12.

⁶ <https://developer.android.com/guide/topics/ui/controls/spinner?hl=pt-br>

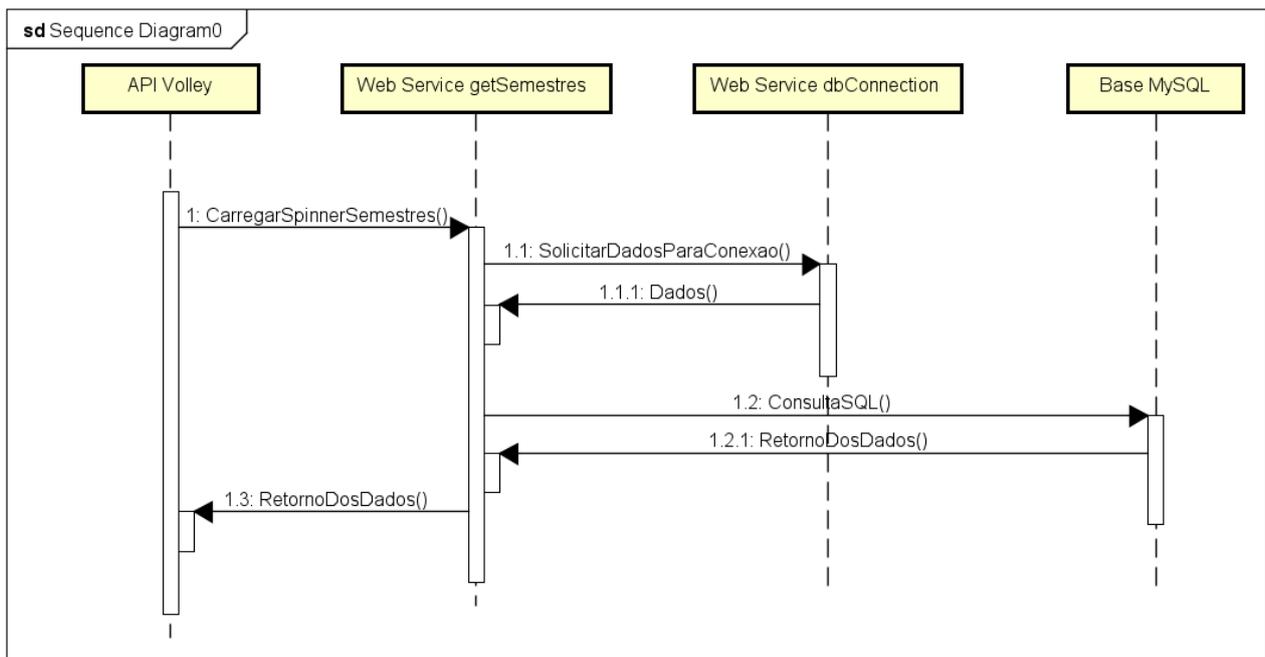


Figura 4.12: Diagrama de Sequência CarregarSpinner

Na operação de Consultar Notas e Faltas, o aluno seleciona o semestre no qual deseja realizar a consulta, a API *Volley* se encarrega de enviar estes dados para o Web Service *getNotas*, que realiza a consulta no banco de dados e retorna o resultado com os dados exibidos na tela, conforme a Figura 4.13.

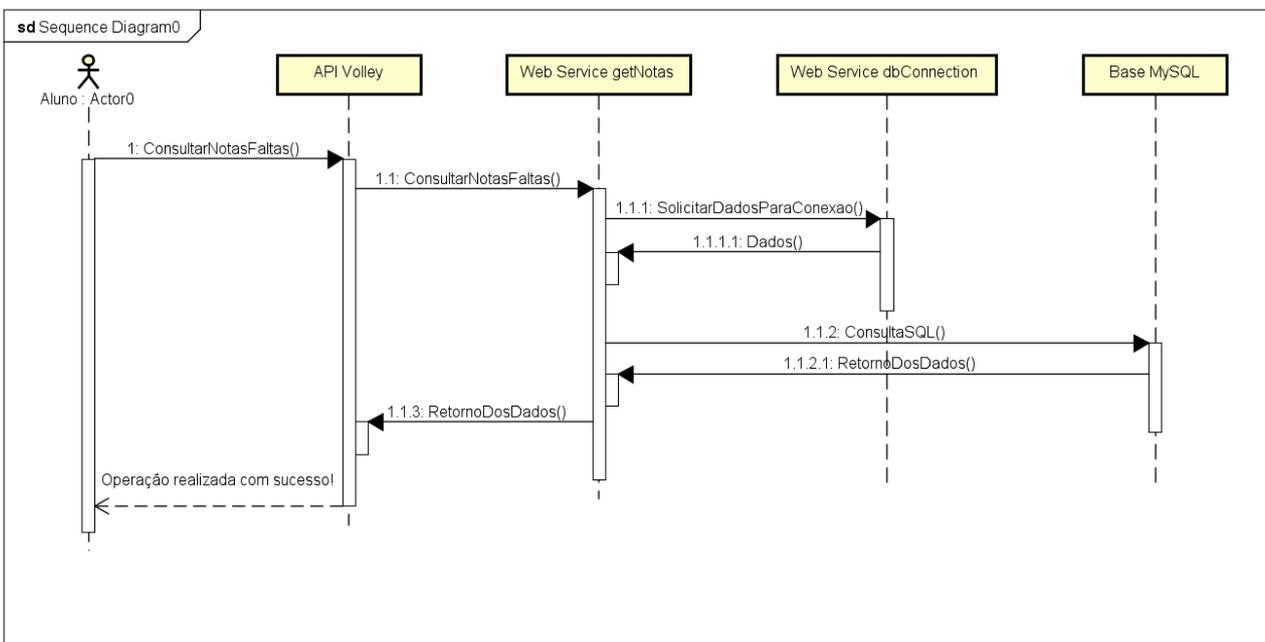


Figura 4.13: Diagrama de Sequência ConsultarNotasFaltas

4.5 IMPLEMENTAÇÃO

A implementação de software é um dos estágios do processo no qual um sistema de software executável é desenvolvido [Sommerville, 2019].

Esta seção, refere-se à implementação das partes iniciais da estrutura do aplicativo PredictApp. Foram implementadas as telas de login e do boletim, esta são exibidas informações de disciplinas cursadas com as suas respectivas notas e faltas. Mais informações, nos links a seguir:

- Protótipo do app na Google Play: <<http://bit.ly/3QzM2pu>>.
- Código-fonte: <<https://github.com/almeidatarcisio/PredictApp>>.
- Video de apresentação: <<https://youtu.be/uQGbYUIId8dYk>>.

A tela inicial é onde o aluno realiza a operação de login, conforme a Figura 4.14.

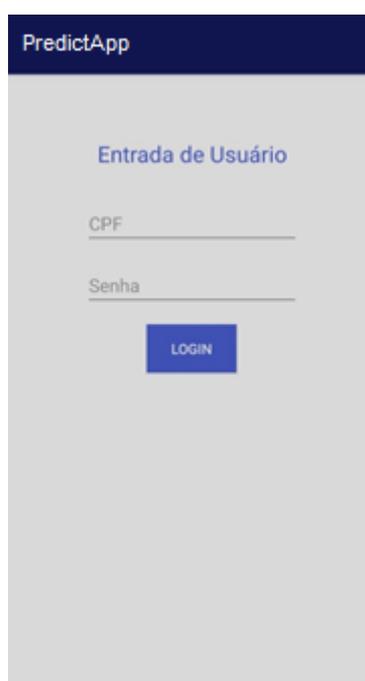


Figura 4.14: Tela login

Nesta tela, o aluno irá digitar seus dados de CPF e senha que serão checados pelo web service via consulta ao banco de dados da instituição de ensino. A classe LoginActivity do Android possui um método chamado validarLogin que realiza a chamada via HTTP ao web service getUsers, localizado no servidor do Heroku. Caso o usuário insira algum dado (CPF ou senha) incorreto, o aplicativo retornará uma mensagem informando.

Ao logar com sucesso, o aplicativo exibirá a próxima tela, onde serão exibidos os dados do aluno (CPF e senha) e os dados relacionados ao boletim do semestre atual (semestre, disciplinas, turma, notas e faltas), conforme Figura 4.15a.

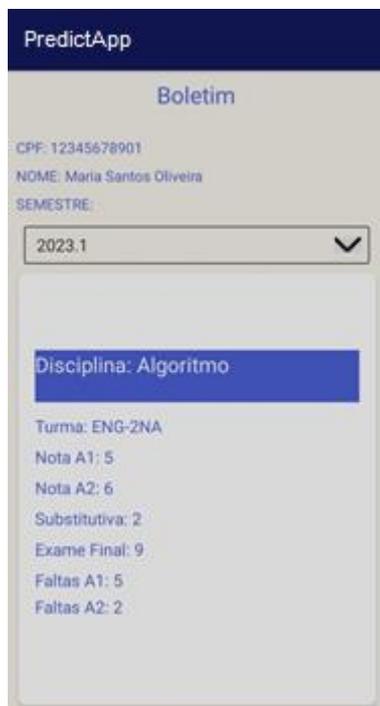


Figura 4.15a: Tela Notas e Faltas



Figura 4.15b: Tela Spinner Semestres

Ao carregar esta tela, a classe NotasActivity do Android possui um método chamado retrieveJSONSemestres que realiza a chamada via HTTP ao web service getSemestres, localizado no servidor do Heroku, que irá retornar os semestres cadastrados na base dados da instituição e carregará o Spinner da figura 4.15a. Em seguida, a mesma classe NotasActivity do Android executa o método retrieveJSONNotas que realiza a chamada via HTTP ao web service getNotas, também localizado no servidor do Heroku, que irá retornar os dados relacionados ao boletim do aluno do semestre atual, conforme a Figura 4.15a.

Ao tocar no Spinner, os semestres cadastrados na base de dados são exibidos na tela do aplicativo, conforme Figura 4.15b. Ao selecionar um semestre, a classe NotasActivity do Android executa novamente o método retrieveJSONNotas, conforme explicado anteriormente, retornando os dados relacionados ao boletim do aluno do semestre selecionado.

4.6 AVALIAÇÃO DE DESEMPENHO

Nesta seção, apresentam-se as métricas adotadas e que são aplicadas para medir o desempenho do aplicativo. A ferramenta Android Profiler foi utilizada para realizar as medições durante o uso do aplicativo.

O Android Profiler é uma ferramenta integrada no Android Studio que permite monitorar o desempenho do aplicativo em tempo real. Ele oferece informações sobre o uso da CPU, memória e outros recursos do dispositivo durante a execução do aplicativo [Android, 2023].

As medições foram realizadas em um smartphone com processador de 1.8Ghz Chipset Snapdragon 632 Qualcomm, GPU Adreno 506, 2GB de memória RAM, 32GB de memória de armazenamento interno, bateria de Litio 3000mAh, sistema operacional Android 10. As métricas adotadas neste trabalho são: Utilização de CPU, Utilização de Memória e Consumo de Energia.

4.6.1 Utilização de CPU

Foi realizado o monitoramento da utilização da CPU pelo aplicativo ao longo do tempo. A ferramenta exibe informações detalhadas sobre os métodos que estão consumindo mais tempo de CPU, ajudando a identificar gargalos de desempenho.

A análise da utilização de CPU é um processo que envolve monitorar e avaliar o uso da CPU para identificar partes do código que estão consumindo mais tempo de processamento. Isso permite identificar gargalos de desempenho e otimizar o aplicativo para garantir uma execução mais eficiente [Android, 2023].

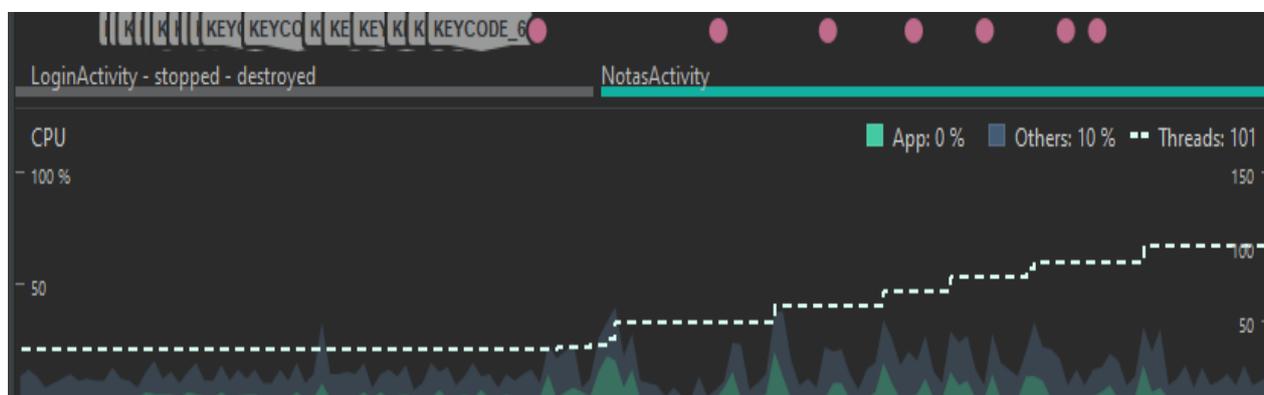


Figura 4.16: Utilização de CPU

Observa-se que o uso de CPU aumenta quando a classe `NotasActivity` é executada, ou seja, no momento em que são realizadas as solicitações aos Web Services e obtidas as respostas às consultas. No entanto não ultrapassa o uso de 50% de CPU, conforme a Figura 4.16.

4.6.2 Utilização de Memória

Foi realizado o monitoramento do uso de memória pelo aplicativo. A ferramenta exibe informações sobre a alocação e liberação de memória, permitindo detectar vazamentos e otimizar o consumo de memória.

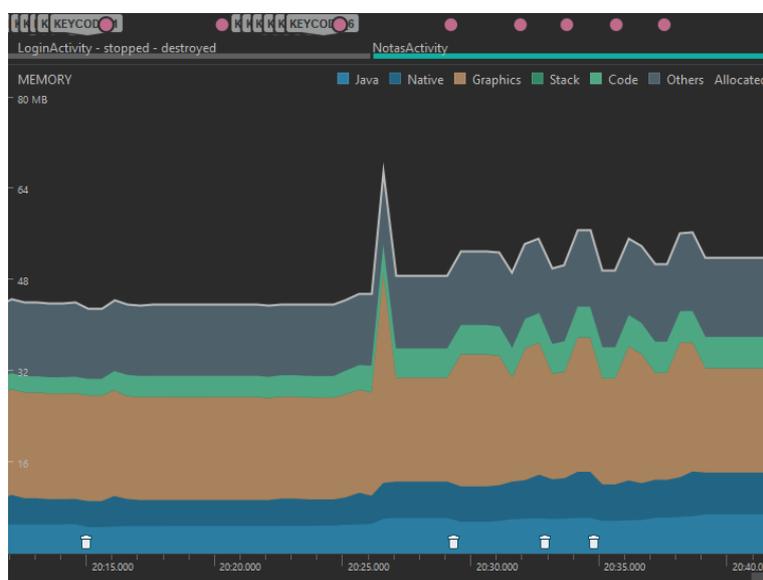


Figura 4.17: Utilização de Memória

Observa-se um pico de uso de memória no intervalo em que a classe `LoginActivity` encerra o seu ciclo de vida e a classe `NotasActivity` passa a ser executada, pois é o momento em que todos os Web Services tem seus serviços solicitados e retornam dados ao aplicativo. O uso de memória supera a marca de 60MB, conforme a Figura 4.17.

A análise da utilização de memória é o processo de monitorar e avaliar o uso de memória para identificar alocações excessivas de memória, vazamentos de memória e outros problemas relacionados à gestão de memória. Isso é crucial para garantir que o aplicativo utilize a memória de maneira eficiente e não cause problemas de desempenho ou travamentos devido a problemas de memória [Android, 2023].

4.6.3 Consumo de Energia

Além dos recursos de monitoramento de CPU e memória, a ferramenta também pode medir o consumo de energia do aplicativo, ajudando a identificar como o uso de recursos está afetando a vida útil da bateria. O gráfico é apresentado na Figura 4.18.

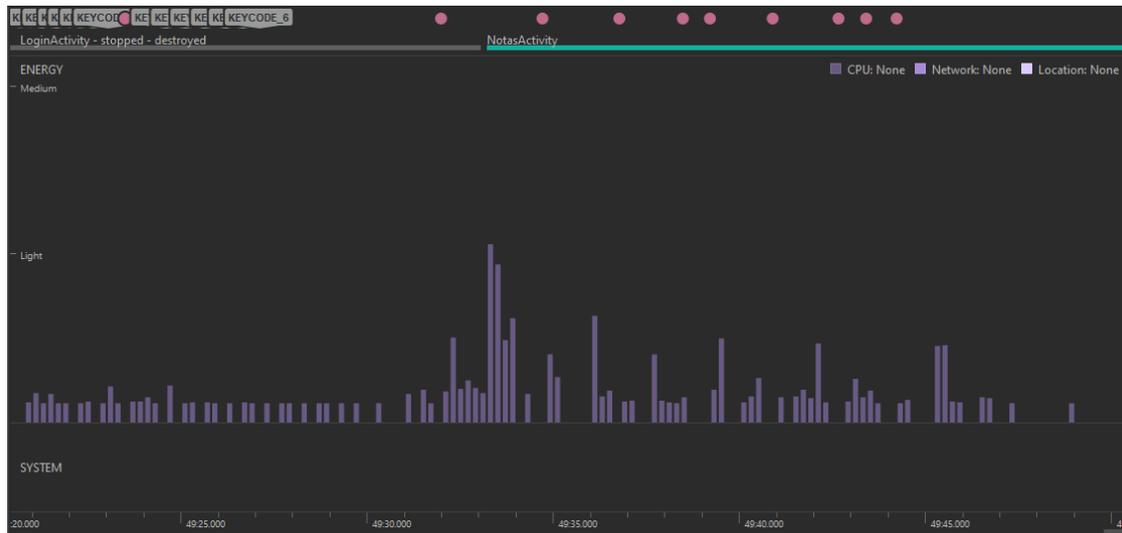


Figura 4.18: Consumo de Energia

A análise do consumo de energia envolve monitorar e avaliar a utilização da energia do dispositivo ao longo do tempo. Isso é importante para garantir que o aplicativo não esteja consumindo energia excessiva, o que pode levar a uma redução na vida útil da bateria e impactar negativamente a experiência do usuário [Android, 2023].

Na análise do gráfico, constata-se que há um pico de consumo de energia no momento da interação com os Web Services, mas dentro de um nível aceitável.

Neste capítulo apresentou-se a arquitetura do sistema com os artefatos (diagramas) gerados nas fases de análise e projeto do sistema, bem como a implementação do protótipo, o *feedback* dos usuários e uma avaliação de desempenho por meio de medições.

5 CONCLUSÕES

A tecnologia Web Services tem se propagado bastante como solução para problemas de integração entre sistemas computacionais. Com a proposta de oferecer serviços distribuídos em uma rede onde podem ser acessados de qualquer lugar, eles chamam atenção dos grandes provedores de tecnologia. Através do uso da arquitetura RESTfull, além da independência de linguagem de programação, a interação entre as implementações dos sistemas torna-se mais fácil.

A interoperabilidade dos Web Services está baseada na implementação de funções simples transportando pouco volume de dados através de arquivos JSON. Essa tecnologia atende ao requisito na realização desse trabalho, facilitando a disponibilização dos serviços mesmo que em linguagens de programação diferentes, sendo um canal mais fácil e barato de compartilhamento de informações, expondo-as de forma padrão e que poderão ser aproveitadas por futuras aplicações.

Durante a Revisão Sistemática da Literatura realizada, ficou evidente que a Mineração de Dados Educacionais é uma área vasta, mas que ainda necessita de maior estruturação e estudo, devido à diversidade e ao volume significativo de dados na área educacional, que é bastante ampla. Contudo, através da pesquisa observa-se que a utilização de dados educacionais traz benefícios significativos para a gestão escolar.

Os artigos selecionados através dos critérios de inclusão, exclusão e qualidade demonstraram conter conteúdo teórico relevante, permitindo uma compreensão aprofundada sobre o tema em questão, sua importância e como ele pode impactar a gestão educacional. No entanto, uma questão de pesquisa (Q3) ainda não foi respondida de maneira abrangente com esta revisão, o que indica a necessidade de estender a pesquisa para outras bases de dados.

Foi possível compreender que os algoritmos determinísticos, como Árvore de Decisão J48 e Naive Bayes, se destacam na literatura como modelos mais eficazes na predição de desempenho profissional. Observou-se também que é passível de automação sistemas com esse tipo de previsibilidade. Além disso, observou-se o uso de vários atributos/variáveis, como idade, gênero, experiência de estágio, notas, instituição, coeficiente de rendimento, habilidades técnicas, comunicação, trabalho em equipe, dentre outras.

Neste trabalho, desenvolveu-se um estudo da utilização da tecnologia Web Services na integração com dispositivos móveis. Foi realizada a análise, projeto e implementação da integração entre um sistema acadêmico e um aplicativo para dispositivos móveis baseado em Web Services, ampliando as possibilidades de o aluno consultar suas notas, frequências e predição de desempenho acadêmico.

Foram utilizadas técnicas de Engenharia de Software em todas as fases do desenvolvimento, compreendendo as principais atividades dos usuários do sistema e seus envolvidos, buscando a melhor forma de abranger o máximo de requisitos. Foi iniciado a implementação de um protótipo de aplicativo para dispositivos móveis para visualização dos resultados. O protótipo do aplicativo desenvolvido tem a intenção de aprimorar o desempenho profissional dos estudantes e com a sua implantação, poderá ser utilizado por todas as partes interessadas. Além disso, foi conduzida uma avaliação de desempenho do protótipo, através da métrica de medição.

Vale ressaltar que se houver uma expansão nas funcionalidades do aplicativo, novos Web Services podem ser desenvolvidos para atender às demandas. Portanto, deve-se observar exatamente em qual contexto esta tecnologia pode e deve ser utilizada, como no caso em estudo, onde ela pode trazer maiores ganhos com a interoperabilidade das aplicações que solicitam apenas poucos recursos de operação.

A experiência vivenciada nesse trabalho possui características próprias que não se pode deixar de registrá-las como prática de lições aprendidas e para aproveitamento em um trabalho futuro. Pode-se destacar:

- A revisão sistemática da literatura, que serviu para reunir, avaliar e sintetizar todas as evidências disponíveis sobre as questões de pesquisa, oferecendo uma visão completa e imparcial do estado do conhecimento atual. Isso é fundamental para a progressão do conhecimento e para a tomada de decisões sobre o tema abordado.
- A elaboração da arquitetura do sistema, que foi um grande passo a ser seguido antes de começar a sua implementação. Entender a estrutura do sistema e o seu comportamento foi fundamental.

Ao longo do desenvolvimento deste trabalho, puderam ser identificadas algumas possibilidades de melhoria e de continuação a partir de futuras pesquisas, as quais incluem:

- A implementação da tela de predição de desempenho acadêmico dos estudantes;
- A possibilidade do uso de diversas técnicas de classificação;
- A integração do sistema acadêmico e do aplicativo para dispositivos móveis com o modelo de classificação para predição de desempenho acadêmico;
- O aplicativo também terá a capacidade de oferecer orientações com base em informações pessoais e acadêmicas fornecidas, a fim de auxiliar na melhoria do desempenho futuro.

O aplicativo criado está atualmente disponível para dispositivos Android, e em breve estará disponível para iOS. Os resultados gerados pelo aplicativo proporcionam uma experiência satisfatória para os usuários, permitindo que os estudantes tomem medidas para aprimorar seus resultados, que são exibidos e, posteriormente, trabalhem para melhorar seu desempenho. Além disso, os professores desempenham um papel fundamental ao inspirar os alunos a alcançarem melhores resultados ao longo do curso e ao orientá-los adequadamente.

REFERÊNCIAS

- AHMED, Saja Taha et al. **Developed third iterative dichotomizer based on feature decisive values for educational data mining**. Indonesian Journal of Electrical Engineering and Computer Science, 18(1):209–217, 2020.
- AMARAL, A. C. **Resolvendo Alguns Conflitos Semânticos em Sistemas de Integração de Dados**. 2005. Trabalho de Graduação em Ciência da Computação, Centro de Informática, UFPE.
- ANDRADE, Rogério Rodrigues; **Estudo de caso – Sistema de gerenciamento de pedidos utilizando web service**. 2008. Monografia (Graduação em Ciência da Computação) – Universidade de Itaúna, Minas Gerais.
- ANDROID. **Android: O Android Profiler**, 2023. Android Studio. Disponível em <https://developer.android.com/studio/profile/android-profiler?hl=pt-br>. Acesso em: 03 de agosto de 2023.
- ARAUJO, Haroldo Alexandre de et al. **Algoritmo simulated annealing: uma nova abordagem**. 2001.
- BOLCH, G. et al. **Queueing Networks and Markov Chains: modeling and performance evaluation with computer science applications**. [S.l.]: Wiley-Interscience, 2006. 896p.
- COSTA, M.; SILVA, R.; BARBOSA, F. **Computação em Nuvem - Parte 1: O que contratar?** INFOQ, 2017. Disponível em < <https://www.infoq.com/br/articles/contratar-computacao-nuvem/>>. Acesso em: 02/08/2023.
- CRESWELL, J.W. **Projeto de pesquisa : métodos qualitativo, quantitativo e misto**. Artmed, Porto Alegre, 3. ed. edition, 2010.
- CRUZ, Maria Elisa Linda Taeza; ENCARNACION, Riah Elcullada. **Analysis and prediction of students' academic performance and employability using data mining techniques: A research travelogue**. The Eurasia Proceedings of Science Technology Engineering and Mathematics, v. 16, p. 117-131, 2021.
- DANIELSSON, Patrik; POSTEMA, Tom; MUNIR, Hussan. **Heroku-Based Innovative Platform for Web-Based Deployment in Product Development at Axis**. IEEE Access 9 (2021): 10805-0819. Web.
- DITTRICH, K. R., ZIEGLER, P. **Three Decades of Data Integration – All Problems Solved?**. 2004.
- FIELDING, R. **Architectural Styles and the Design of Network-based Software Architectures**, PhD Dissertation, University of California, Irvine, California, USA, 2000.

GULERIA, Pratiyush; SOOD, Manu. **Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling**. Education and Information Technologies, v. 28, n. 1, p. 1081-1116, 2023.

HEROKU. **Heroku: Cloud Application Platform**, 2023. Salesforce Developers. Disponível em <https://www.heroku.com/>. Acesso em: 27 de julho de 2023.

HU, Qian; RANGWALA, Huzefa. **Towards fair educational data mining: A case study on detecting atrisk students**. In Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020), pages 431–437, 2020.

HUNG, Hui-Chun et al. **Applying educational data mining to explore students' learning patterns in the flipped learning approach for coding education**. Symmetry, 12(2):213, 2020.

HUSSAIN, Sadiq et al. **Educational data mining and analysis of students' academic performance using weka**. Indonesian Journal of Electrical Engineering and Computer Science, 9(2):447–459, 2018.

JASSIM, Mustafa Abdalrassual. **Analysis of the performance of the main algorithms for educational data mining: a review**. In: IOP conference series: materials science and engineering. IOP Publishing, 2021. p. 012084.

KAUR, Amritpreet et al. **Classification and prediction based enhanced j48 and reptime algorithms to predict corona virus pandemic**. no. July, 2020.

KEELE, Staffs et al. **Guidelines for performing systematic literature reviews in software engineering**. 2007.

KITCHENHAM, Barbara A. **Systematic review in software engineering: where we are and where we should be going**. In: Proceedings of the 2nd international workshop on Evidential assessment of software technologies. 2012. p. 1-2.

KUH, George D. et al. **What matters to student success: A review of the literature**. Washington, DC: National Postsecondary Education Cooperative, 2006.

LECHETA, Ricardo. **Google Android: Aprenda a Criar Aplicações para Dispositivos Móveis com Android SDK**. 5ª Edição. São Paulo: Novatec, 2015.

LECHETA, Ricardo. **Web Services RESTful: Aprenda a criar web services RESTful em Java na nuvem do Google**. São Paulo: Novatec, 2015.

LEE, L. E. et al. **Evaluation of prediction algorithms in the student dropout problem**. Journal of Computer and Communications, 8(03):20, 2020.

LILJA, D. J. **Measuring Computer Performance: a practitioner's guide**. [S.l.]: Cambridge University Press, 2005. 278p

MACHIDA, F.; ANDRADE, E.; KIM, D. S.; TRIVEDI, K. S. **Candy: Component-based availability modeling framework for cloud service management using sysml**. In: IEEE. 2011 IEEE 30th International Symposium on Reliable Distributed Systems. [S.l.], 2011. p. 209–218.

- MAHAJAN, Ginika; SAINI, **Bhavna**. **Educational data mining**: A state-of-the-art survey on tools and techniques used in edm. *International Journal of Computer Applications & Information Technology*, 12(1):310–316, 2020.
- MAHESHWARI, Ekansh et al. **Prediction of factors associated with the dropout rates of primary to high school students in india using data mining tools**. In *Frontiers in Intelligent Computing: Theory and Applications*, pages 242– 251. Springer, 2020.
- MAHMOUD, Ousay H. **Service-Oriented Architecture (SOA) and Web Services**: The Road to Enterprise Application Integration (EAI). 2005. Disponível via URL em: <http://java.sun.com/developer/technicalArticles/WebServices/soa/>. Acesso em 30/12/2008.
- MARQUES, L.T., MARQUES, B.T., SILVA, C.A.M. . **A Descoberta das Causas da Retenção Acadêmica Utilizando Mineração de Dados**: Uma Revisão Sistemática da Literatura. *RENOTE*, Porto Alegre, v. 20, n. 1, p. 263–272, 2022.
- MASCI, Chiara; JOHNES, Geraint; AGASISTI, Tommaso. **Student and school performance across countries**: A machine learning approach. *European Journal of Operational Research*, v. 269, n. 3, p. 1072-1085, 2018.
- MENASCÉ, D. A.; ALMEIDA, V. A. F. **Performance by Design**: computer capacity planning by example. [S.l.]: Prentice Hall PTR, 2005. 462p.
- MENIYA, A.D.; JETHVA, H.B. **Next Generation Mobile Application in Cloud Computing using RESTful Web Services**. *Int.J.Computer Technology & Applications (www.ijcta.com)*, Vol 3 (3), 949-952, 2012.
- NASCIMENTO, F.A. et al. **MINERAÇÃO DE DADOS EDUCACIONAIS: UMA REVISÃO SISTEMÁTICA DA LITERATURA**. Rio de Janeiro: Editora Oston Ltda. 2023. DOI: 10.5281/zenodo.7763620
- ORT, Ed. **Service-Oriented Architecture and Web Services: Concepts, technologies, and tools**. 2005. Disponível via URL em: <http://java.sun.com/developer/technicalArticles/WebServices/soa2/>. Acesso em 30/12/2008.
- PEIXOTO, M. L. M. **Oferecimento de QoS para computação em nuvens por meio de metaescalonamento**. Tese (Doutorado) — Universidade de São Paulo, 2012.
- PRESSMAN, Roger S. **Engenharia de Software**. 8ª edição. São Paulo: McGrawHill, 2016.
- QIU, Jiezhong et al. **Gcc: Graph contrastive coding for graph neural network pre-training**. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1150–1160, 2020.
- REPASO, Jennifer Anne A.; CAPARIÑO, Elenita T. **Analyzing and predicting career specialization using classification techniques**. *International Journal*, v. 9, n. 1.3, 2020.
- REZENDE, Solange Oliveira. **Sistemas inteligentes**: fundamentos e aplicações. Editora Manole Ltda, 2003.

- SANTOS, Alfredo Luiz dos. **Integração de Sistemas com Java**. Rio de Janeiro: Brasport, 2007.
- SOMMERVILLE, I. **Engenharia de Software**. 10ª edição. São Paulo: Pearson, 2019.
- SOUZA, V.F., SANTOS, T.C.B. **Processo de Mineração de Dados Educacionais aplicado na Predição do Desempenho de Alunos**: Uma comparação entre as Técnicas de Aprendizagem de Máquina e Aprendizagem Profunda. Revista Brasileira de Informática na Educação – RBIE, 2021.
- TAEZA-CRUZ, Maria Elisa Linda; CAPILI-KUMMER, Marifel Grace. **Decision Support System to Enhance Students' Employability using Data Mining Techniques** for Higher Education Institutions. International Journal of Computing and Digital Systems, 2023.
- VIHERÄKOSKI, Johanna. **Strengths in career development**: Modeling strength-based career counseling through reflecting customer experience. 2020.
- VIVIAN, R. L. et al. **Mineração de Dados Educacionais e Análise de Sentimentos em Ambientes Virtuais de Aprendizagem**: um Mapeamento Sistemático. EaD em Foco, v. 12, n. 2, e1786, 2022.
- WAHEED, Hajra et al. **Predicting academic performance of students from VLE big data using deep learning models**. Computers in Human behavior, v. 104, p. 106189, 2020.
- YANG, Cuibi; HUAN, Shuliang; YANG, Yong. **A practical teaching mode for colleges supported by artificial intelligence**. International Journal of Emerging Technologies in Learning (IJET), v. 15, n. 17, p. 195-206, 2020.
- YORK, Travis T.; GIBSON, Charles; RANKIN, Susan. **Defining and measuring academic success. Practical assessment, research, and evaluation**, v. 20, n. 1, p. 5, 2015.
- YULIANTO, Lili Dwi; TRIAYUDI, Agung; SHOLIHATI, Ira Diana. **Implementation Educational Data Mining For Analysis of Student Performance Prediction with Comparison of K-Nearest Neighbor Data Mining Method and Decision Tree C4. 5**: Implementation Educational Data Mining For Analysis of Student Performance Prediction with Comparison of K-Nearest Neighbor Data Mining Method and Decision Tree C4. 5. Jurnal Mantik, v. 4, n. 1, p. 441-451, 2020.
- ZOTA, R.; PETRE, I. A. **An overview of the most important reference architectures for cloud computing**. Informatica Economica, v. 18, n. 4, 2014

APÊNDICE A – FORMULÁRIO DE PERGUNTAS

NOME DO ENTREVISTADO:

FUNÇÃO:

1. Atualmente, como o aluno obtém suas notas das disciplinas e uma predição de desempenho acadêmico?
2. Quais são os problemas no acesso ao portal acadêmico e como um novo sistema ajudaria a diminuir esses problemas?
3. As informações do portal acadêmico são transferidas para outros sistemas dentro da própria instituição e também podem ser recebidas destes?
4. Quais as suas expectativas quanto à criação de um aplicativo para consulta de notas, frequências e predição de desempenho acadêmico?
5. Que contribuição direta a integração dos sistemas trará para os objetivos da instituição de ensino?
6. Anteriormente, houve alguma tentativa de utilização de algum tipo de tecnologia para o problema?
7. O que precisa e o que não precisa ser compatível com o app?
8. Quais suas sugestões a nível técnico sobre a integração dos sistemas?

Observações: