



Universidade Estadual de Feira de Santana
Programa de Pós-Graduação em Computação Aplicada

Análise da relação entre os fatores interno e externo e o desempenho acadêmico dos estudantes do curso de Engenharia da Computação da UEFS

Naan Silva Cardoso

Feira de Santana

2019



Universidade Estadual de Feira de Santana
Programa de Pós-Graduação em Computação Aplicada

Naan Silva Cardoso

Análise da relação entre os fatores interno e externo e o desempenho acadêmico dos estudantes do curso de Engenharia da Computação da UEFS

Dissertação apresentada à Universidade Estadual de Feira de Santana como parte dos requisitos para a obtenção do título de Mestre em Computação Aplicada.

Orientador: João B. da Rocha-Junior

Coorientadora: Claudia Pinto Pereira

Feira de Santana

2019

Ficha Catalográfica - Biblioteca Central Julieta Carteado - UEFS

C265

Cardoso, Naan Silva

Análise da relação entre os fatores interno e externo e o desempenho acadêmico dos estudantes do curso de Engenharia da Computação da UEFS / Naan Silva Cardoso . – 2019.

71 f.: il.

Orientador: João Batista da Rocha-Junior.

Coorientadora: Claudia Pinto Pereira.

Dissertação (mestrado) – Universidade Estadual de Feira de Santana, Programa de Pós-graduação em Computação Aplicada, Feira de Santana, 2019.

1. Estudantes – desempenho acadêmico. 2. Curso de Engenharia da Computação – UEFS. 3. Participação. 4. Mídias sociais. 5. Mineração de dados. I. Rocha-Junior, João Batista da, orient. II. Pereira, Claudia Pinto, coorient. III. Universidade Estadual de Feira de Santana. IV. Título.

CDU: 004-057.875:378(814.22)

Luis Ricardo Andrade da Silva - Bibliotecário - CRB-5/1790

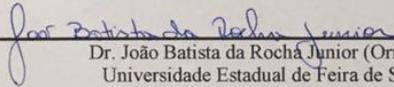
Naan Silva Cardoso

Análise da relação entre os fatores interno e externo e o desempenho acadêmico dos estudantes do curso de Engenharia da Computação da UEFS

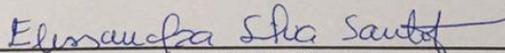
Dissertação apresentada à Universidade Estadual de Feira de Santana como parte dos requisitos para a obtenção do título de Mestre em Computação Aplicada.

Feira de Santana, 30 de agosto de 2019

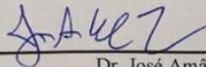
BANCA EXAMINADORA



Dr. João Batista da Rocha Junior (Orientador)
Universidade Estadual de Feira de Santana



Dra. Elissandra Silva Santos
Faculdade do Nordeste da Bahia



Dr. José Amâncio Santos
Universidade Estadual de Feira de Santana

Abstract

Academic performance, with respect to the student concept (i.e. grades, evaluation, approval/disapproval), has been one of the challenges that the Brazilian educational system has been facing in recent years. Some indicators show that the retention and dropout rates are high in the Computing courses. In this research, we sought to analyze the hypothesis that there is a relationship between internal (participation) and external (use of Social Media) factors and student academic performance. For this purpose, data collected from the smartphones of the subjects of this research and of the observation technique within the classroom were analyzed in order to confirm or refute such hypothesis. Data mining techniques such as correlation analysis are used to identify relationship between variables. The research was carried out with the students from Algorithms and Programming I, Algorithms and Programming II and two Data Structure classes T01 and T02, in the Computer Engineering course at the State University of Feira de Santana - UEFS. The results of the study indicate that for the internal factor, no statistically significant result was obtained to confirm the hypothesis that there is a relationship between student participation in classes and academic performance. For the external factor, the sample data points to a strong positive relationship between the use of Facebook in the performance and a significant negative relation between the use of Instagram, confirming the hypothesis that there is a relation between the external factor (Social Media) and the academic performance.

Keywords: Higher Education, Academic Performance, Data Mining, Participation, Media Networks.

Resumo

O desempenho acadêmico, no que diz respeito ao conceito do estudante (i.e. notas, avaliação, aprovação/reprovação), tem sido um dos desafios que o sistema educacional brasileiro vem enfrentando nos últimos anos. Alguns indicadores mostram que os índices de retenção e evasão são altos nos cursos da área de Computação. Nesta pesquisa, procurou-se analisar a hipótese de que existe relação entre os fatores interno (participação) e externo (uso de Mídias Sociais) e o desempenho acadêmico do estudante. Para isso, dados coletados dos *smartphones* dos sujeitos dessa pesquisa e da técnica de observação dentro da sala de aula foram analisados para confirmar ou refutar tais hipóteses. Técnicas de mineração de dados, como análise de correlação são utilizadas para identificar relação entre as variáveis. A pesquisa foi realizada com os estudantes das turmas de Algoritmos e Programação I, Algoritmos e Programação II e duas turmas de Estrutura de Dados T01 e T02, no curso de Engenharia da Computação na Universidade Estadual de Feira de Santana. Os resultados do estudo indicam que, para o fator interno, nenhum resultado estatisticamente significativo foi obtido para confirmar a hipótese de que existe relação entre a participação dos estudantes nas aulas e o desempenho acadêmico. Para o fator externo, os dados da amostra apontam para uma relação forte positiva significativa entre o uso do (*Facebook*) e o desempenho e uma relação negativa significativa entre o uso do (*Instagram*), confirmando a hipótese de que existe relação entre o fator externo (Mídias Sociais) e o desempenho acadêmico.

Palavras-chave: Ensino Superior, Desempenho Acadêmico, Mineração de Dados, Participação, Mídias Sociais.

Prefácio

Esta dissertação de mestrado foi submetida a Universidade Estadual de Feira de Santana (UEFS) como requisito parcial para obtenção do grau de Mestre em Computação Aplicada.

A dissertação foi desenvolvido dentro do Programa de Pós-Graduação em Computação Aplicada (PGCA) tendo como orientador o Dr. **João B. da Rocha-Junior** e coorientadora a Dra. **Claudia Pinto Pereira**.

Agradecimentos

Pensando eu que essa seria a parte mais fácil, enganei-me!

Vou começar pela melhor parte que é agradecendo a Deus pelo dom da vida, sem Ele eu não estaria aqui. Mas, também quero agradecê-lo principalmente por acreditar em mim e fazer com que eu acredite e não perca a fé em mim nunca. Obrigada, Deus!

À minha família, obrigada sempre, sem vocês eu perco o chão. À minha mãe Rosângela, mulher de fibra e de um coração gigante, meu exemplo de vida, obrigada por sempre lutar por nós, te amo. Aos meus irmãos Daniel e Kassyla, somos sempre um por todos e todos por um, e aos meus sobrinhos Lara, Daniel Filho e Cecília pelos momentos de alegria e diversão quando estamos juntos, e Lorenzo que ainda nem chegou mas já mandou um “tchauzinho” pelo monitor para a titia. Somos uma linda família.

Aos meus familiares. Minha avó, que sempre lutou e torceu por nós, a senhora é e sempre foi guerreira, te amo! Aos meus tios, tias, primos e primas, lembro que Tia Rose e Tia Vana sempre me diziam: *“está acabando, tenha calma”*, é acabou!! E minha prima Karol que em uma de nossas longas conversas falou: *“Ná, te vendo no mestrado fico com vontade de fazer também, mas quando te vejo correndo com tanta coisa pra fazer, desanima”*. Desanima não, Rol, no final escrever os agradecimentos até que é divertido! :)

Aos amigos. Não tenho tantos amigos assim, mas os que eu tenho são verdadeiros e sempre que possível nos encontramos e a diversão é garantida. Dedico esta frase a vocês, meus amigos, mas não vou citar nomes para não correr o risco de esquecer algum e apanhar :) *“...A vida é isso: pessoas vão e vem, não tem jeito. E a gente vai selecionando quem fica ou não no lado esquerdo do peito”* (Autor desconhecido).

Meu orientador. João, obrigada por tudo, os ensinamentos, o apoio, por ter me aceitado como orientanda (na verdade fui roubada rs), pela paciência e, acima de tudo por não ter permitido que eu interrompesse o processo quando achei que não dava mais para continuar. À minha coorientadora, Claudia. More, obrigada acima de tudo pela amizade, pela competência, os conselhos, a insistência, as brigas (essa foi a pior parte), as cobranças, por não me deixar fracassar e por não desistir de mim. Deus escolhe as pessoas certas para colocar em nosso caminho. Aqui lhes exprimo a minha eterna gratidão.

Aos professores que participaram da minha banca de defesa, Elissandra e Amâncio. Muito obrigada pelas contribuições e os elogios (esses foram muitos). Tenham certeza de que as sugestões de vocês ajudaram e muito com a melhoria e o enriquecimento deste trabalho.

Uau, de repente agradecer começou a ficar fácil :)

Aos colegas do mestrado. Ayala e Jefferson, nossa caminhada juntos vem desde a graduação, olha só onde chegamos?! À Felipe Pains, Ricardo de Jesus, Luiz Gustavo, Fabrício e Emerson, lembram das nossas horas, dias e noites estudando estrutura e banco de dados? Dizíamos que depois destas disciplinas nada mais nos abalava. É isso, conseguimos e hoje somos mais que colegas, somos amigos. Do mestrado para a vida!!! uhuuuu

Por fim e não menos importante, os agradecimentos são igualmente devidos aos colegas Claudio Valiense e Gabriel Antônio. Obrigada pelo apoio e a colaboração no desenvolvimento do aplicativo *Cookie*. Daniel Andrade, “Dan” obrigada por tudo, sua colaboração e apoio foram vitais no momento de processar os dados, *you are the best!* :). Também não posso esquecer dos estudantes das disciplinas de Algoritmos I, Algoritmos II e as duas turmas de Estrutura de Dados do curso de Engenharia da Computação da UEFS, que ajudaram participando da pesquisa. Foi um semestre divertido, nos tornamos parceiros, lembro de quando alguns me paravam na universidade para perguntar: “...e aí, já analisou os dados?”, “depois quero ver o resultado...”. Então, meus queridos, aqui está o fruto colhido com o apoio de vocês, e eu só tenho a dizer, meu muito Obrigada! :)

Limitações são fronteiras criadas apenas pela nossa mente.

Provérbio Chinês

Sumário

Abstract	i
Resumo	ii
Prefácio	iii
Agradecimentos	iv
Sumário	vii
Lista de Tabelas	viii
Lista de Figuras	ix
Lista de Abreviações	x
1 Introdução	1
1.1 Hipóteses	3
1.2 Limites da Pesquisa	4
1.3 Objetivos	5
1.4 Organização do Trabalho	5
2 Fundamentação Teórica	7
2.1 Ensino Superior Brasileiro	7
2.1.1 Retenção e Evasão no Curso de Computação	9
2.1.2 Desempenho Acadêmico	10
2.2 Mineração de Dados	12
2.3 Mineração de Dados Educacionais	14
2.3.1 Tarefas da Mineração de Dados Educacionais	15
2.4 Trabalhos Relacionados	19
3 Metodologia	22
3.1 Observação	23
3.2 Aplicativo Cookie	25
3.3 Comitê de Ética e Pesquisa	28

3.4	Cenário de Pesquisa	29
3.5	Apresentação para as Turmas	31
3.6	Análise e Processamento dos Dados	31
4	Coleta e Análise de Dados	33
4.1	Coleta dos Dados	33
4.2	Aplicação da Mineração de Dados	33
5	Resultados	37
5.1	Estudo experimental dos fatores internos	37
5.2	Estudo experimental dos fatores externos	40
6	Considerações Finais	45
6.1	Pesquisas Futuras	46
	Referências	48
A	Termo de Consentimento Livre e Esclarecido - TCLE (Docente)	54
B	Termo de Consentimento Livre e Esclarecido - TCLE (Discente)	56

Lista de Tabelas

1.1	Informações da Amostra	4
3.1	Variáveis coletadas através da Observação	23
3.2	Variáveis coletadas através do <i>Smartphone</i>	25
4.1	Junção das variáveis coletadas	34
4.2	Variáveis selecionadas na Observação	35
4.3	Variáveis selecionadas do <i>Cookie</i>	36
5.1	Distribuição dos estudantes nas disciplinas em relação ao nível de desempenho acadêmico	38
5.2	Distribuição do desempenho dos estudantes no semestre	38
5.3	Distribuição da participação dos estudantes	39
5.4	Distribuição das turmas em relação ao desempenho acadêmico nas disciplinas.	41
5.5	Distribuição do desempenho dos estudantes no semestre.	42
5.6	Distribuição da utilização dos aplicativos.	42

Lista de Figuras

2.1	Matrículas no Ensino Superior (2007-2017).	8
2.2	Etapas da extração de conhecimentos de dados.	13
2.3	Ciclo de aplicação de mineração de dados em sistemas educacionais.	15
3.1	Esquema para execução do projeto.	23
3.2	Mapa de assentos.	24
3.3	Arquitetura do <i>Cookie</i>	27

Lista de Abreviações

Abreviação	Descrição
ADaM	Advanced Data Management Reserch Group
AVA	Ambiente Virtual de Aprendizagem
BD	Banco de Dados
CEP	Comitê de Ética em Pesquisa
DER	Diagrama Entidade Relacionamento
EaD	Educação à Distância
EDM	Mineração de Dados Educacionais (Education Data Mining)
EMA	Ecological Momentary Assessment
IES	Instituição de Ensino Superior
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
KDD	Knowlegde Discovery from Data
LABOTEC	Laboratório de Tecnologia
MD	Mineração de Dados
MEC	Ministério da Educação
MIT	Instituto de Tecnologia de Massachusetts (Massachusetts Institute of Technology)
PBL	Aprendizado Baseado em Problemas (Problem-Based Learning)
SNA	Análise de Redes Sociais (<i>Social Network Analysis</i>)
TCLE	Termo de Consentimento Livre e Esclarecido
TIC	Tecnologia da Informação e Comunicação
UEFS	Universidade Estadual de Feira de Santana

Capítulo 1

Introdução

“All victories hide an abdication”

–Simone de Beauvoir

Segundo Castells [Castells 2005], desde a chegada do século XX o mundo vem passando por um processo de transformação estrutural devido a velocidade no surgimento e adoção de novas tecnologias. Estas novas tecnologias associadas à internet, “cibercultura” [Levy 2010, p.147], estão cada vez mais frequentes na sociedade, seja por meio de computadores portáteis, *smartphones*, *tablets*, entre outros.

Segundo de Souza e Coelho [de Souza e Coelho 2018] o uso dos *smartphones*, por exemplo, permite que as pessoas permaneçam conectadas a todo momento e de qualquer lugar e, quanto maior é o seu uso, maior é o seu envolvimento e sociabilidade. Esta afirmação é sustentada por Castells, quando diz que “a sociedade em rede é uma sociedade hipersocial, não uma sociedade de isolamento” [Castells 2005, p.23].

O uso constante das tecnologias digitais, bem como de mídias sociais (e.g. *Facebook*, *WhatsApp*, *Instagram*), e o modo como elas estão sendo utilizadas, provocam transformações no comportamento da sociedade que podem ser percebidas na maneira como as pessoas se comunicam e se relacionam. Para [Lévy 1999, Lemos e Lévy 2010], este cenário apresenta um novo meio de comunicação que pode permitir que as pessoas se expressem sem ter a necessidade do poder de um mediador.

Nesta cibercultura, surge uma nova juventude a qual vive conectada à rede, receptiva a novas amizades e compartilhamento de contextos e ideais nas mídias sociais. O jovem atual é capaz de efetuar atividades diversas em paralelo (e.g. estudar e conversar com o colega no *WhatsApp*), e ainda assim assimilar o conteúdo estudado, e este comportamento não é diferente no Ensino Superior. De acordo com de Souza e Coelho [de Souza e Coelho 2018] “em um mundo em que a tecnologia digital evolui

em uma enorme velocidade e proporciona diferentes campos, a educação não pode ficar de fora de sua área de influência” [de Souza e Coelho 2018, p.71].

As instituições precisam estar atentas às necessidades dos estudantes, com relação a utilização de recursos que envolvem o mundo digital, a fim de promover e tornar sua utilização colaborativa nas propostas de ensino [de Souza e Coelho 2018]. O uso e a inclusão destes recursos no sistema educacional podem ajudar no incentivo de melhorias que possam ter relação com o desempenho acadêmico dos estudantes. De acordo com Albuquerque [Albuquerque 2016], o desempenho pode ser constituído por um conjunto de fatores possíveis, tais como, sistema didático-pedagógico, socioeconômicas, cultura, qualidade do ensino, entre outros, o desempenho acadêmico tem sido um fator preocupante para as instituições, professores e alunos.

Conforme os autores Rodrigues e Outros [Rodrigues et al. 2017], o desempenho acadêmico, dentre outras compreensões possíveis, pode ser considerado como o meio de obter resultado positivo ou negativo, através de uma avaliação acadêmica, e que é habitualmente mensurado em termos numéricos de zero a dez. Dessa forma, Cerqueira e Outros afirmam que a universidade deve estar “atenta às dificuldades enfrentadas pelos alunos, para que possa buscar formas de superá-las, garantindo a eficácia do ensino ministrado e a consequente aprendizagem do aluno” [Cerqueira et al. 2000, p.7].

Diante desse contexto, vários estudos vêm sendo feitos com o intuito de encontrar fatores que podem ter relação com o desempenho. As pesquisas realizadas por Miranda e Outros [Miranda et al. 2015], buscam encontrar correlação entre o tempo gasto dedicado aos estudos e o desempenho, Hoed [Hoed 2016] aponta que a interação entre professor-aluno contribui para um bom rendimento.

Nesta pesquisa, a coleta dos dados foi feita de duas maneiras: a observação em sala de aula, pela autora, e a utilização do *Cookie*, um aplicativo desenvolvido para *smartphone*. O *Cookie* é uma ferramenta cujo objetivo principal é coletar os dados dos alunos para conhecer o comportamento do estudante na cibercultura e tentar encontrar relação entre o seu comportamento e o desempenho acadêmico.

Neste trabalho, o desempenho acadêmico do estudante pode estar relacionado a dois grupos de fatores: i) interno: são aqueles os quais estão direcionados ao ambiente vivido dentro da sala de aula e; ii) externo: diz respeito aos fatores relacionados ao ambiente fora da sala de aula. Para isso, foram elaboradas hipóteses, como pode ser visto na Seção 1.1, com a finalidade de confirmar a existência de correlações entre os fatores internos e externos e o desempenho acadêmico dos estudantes.

Os fatores internos foram coletados por meio da técnica de observação. As variáveis é da nossa autoria projeto caracterizadas como fatores internos são: hora de chegada e saída do aluno na aula, hora de chegada e saída intermediária do aluno na sala, local que o aluno sentou, número de participação na aula, hora inicial e final da aula, recursos utilizados, assunto.

Os fatores externos foram obtidos através do aplicativo *Cookie*, instalado no celular dos estudantes. Os dados coletados pelo aplicativo são: a latitude e a longitude, a luminosidade do celular, o nome do aplicativo em uso pelo estudante, o estado do celular (parado ou em movimento) e a data da coleta. O *Cookie* é um aplicativo Android, que foi desenvolvido para ser utilizado nesta pesquisa e tem a finalidade de coletar dados dos estudantes, sem que os mesmos interajam com a ferramenta.

As variáveis coletadas na observação e no *Cookie* foram escolhidas pela autoria deste trabalho, durante a concepção inicial do projeto. O propósito foi de coletar o máximo de variáveis possíveis, e na etapa de seleção e processamento dos dados, escolher aquelas as quais estariam mais voltadas para alcançar o objetivo do trabalho.

Este trabalho utiliza a técnica de mineração de correlação para encontrar relação entre os fatores (interno e externo) e o desempenho acadêmico, utilizando os dados coletados dos estudantes. A pesquisa foi executada com estudantes das disciplinas do primeiro e segundo períodos, voltadas para programação, do curso de Engenharia da Computação, na Universidade Estadual de Feira de Santana - UEFS.

Com a propagação do uso das TICs, impulsionada pelo crescimento acelerado da Internet, um grande volume de dados tem sido gerado e armazenado em depósitos de dados digitais, destinados a fins diversos. Áreas distintas como saúde, comércio, indústria, educação, economia, telecomunicações, governo, entre outras geram uma vasta quantidade de dados, e trabalhos constantes vêm sendo realizados em busca de análises desses dados gerados, uma vez que novos padrões e conhecimentos podem ser identificados.

1.1 Hipóteses

Para a análise deste trabalho, foi escolhida como fator interno, a variável *participação*. Esta variável foi selecionada, pois é considerada como um fator de importância para a aprendizagem do estudante, uma vez que envolve dentre outros, o envolvimento nas atividades e nas aulas e a dedicação aos estudos, além disso, pode influenciar no seu desempenho [Wolf-Wendel et al. 2009], [Porto e Gonçalves 2017], [Moleta et al. 2017] e [Fior e Mercuri 2018].

Já como fator externo, a variável selecionada foi *nome do aplicativo aberto*. Nesta variável, foram escolhidas as mídias sociais, pois o seu uso têm crescido cada vez mais. Dados do *Global Digital Report*¹ mostram que 120 milhões de pessoas acessaram as Mídias Sociais (YouTube, *Facebook*, *WhatsApp* e *Instagram*) através do *smartphone*, em 2017. Outro fator importante é o crescente uso dessas redes no meio acadêmico, pelas facilidades e potencialidades que elas apresentam para o processo de ensino [Souza e Schneider 2014].

¹<https://www.slideshare.net/wearesocial/digital-in-2018-in-southern-america-part-1-north-86863727>

Diante do contexto apresentado, percebe-se a importância em analisar a relação entre os fatores interno e externo e o desempenho acadêmico do estudante. Para isso, as seguintes hipóteses serão analisadas no decorrer deste trabalho:

- Existe uma relação entre o fator interno, a participação do estudante na aula e o desempenho acadêmico.
- Existe uma relação entre o fator externo, o uso das Mídias Sociais (*WhatsApp, YouTube, Instagram e Facebook*) e o desempenho acadêmico de estudantes.

1.2 Limites da Pesquisa

O presente trabalho delimitou-se em coletar dados dos alunos de três disciplinas (Algoritmos e Programação I, Algoritmos e Programação II e Estrutura de Dados) do curso de Engenharia da Computação da UEFS. Pretende-se analisar dados coletados dos estudantes matriculados apenas nas disciplinas Algoritmos e Programação I, 1º semestre, Algoritmos e Programação II, 2º semestre, Estrutura de Dados (Turmas 01 e 02), 2º semestre, mencionadas na Metodologia Capítulo 3.

Em relação às variáveis, que estão focadas no aluno, e na aula, encontram-se detalhadas na Metodologia Capítulo 3.

O curso de Engenharia da Computação da UEFS é ofertado desde 2003. Atualmente, possui 360 alunos ativos, sendo que nas turmas selecionadas para a análise constam, inicialmente, 96 alunos no total, distribuídos conforme Tabela 1.1.

Tabela 1.1: Informações da Amostra

Disciplinas	Nº de alunos matriculados	Carga Horária
Algoritmos e Programação I	24	60
Algoritmos e Programação II	24	30
Estrutura de Dados (Turma 01)	24	30
Estrutura de Dados (Turma 02)	24	30

Fonte: Elaborado pelo Próprio Autor a partir de dados da UEFS.

Como pode ser observado na Tabela 1.1, a disciplina de Estrutura de Dados é dividida em duas turmas (T01 e T02).

Informações como quantidade de livros locados na biblioteca, bolsa estudantil, titulação e número de professores no curso, avaliação institucional e docente, entre outros, fazem parte do banco de dados da Instituição de Ensino Superior - IES. Entretanto, estes fatores socioeconômicos e culturais não são considerados na análise, pois o foco deste trabalho é analisar apenas as variáveis coletadas com o uso da observação em aula e o uso do aplicativo pelos estudantes.

As tomadas de decisões advindas desta pesquisa a partir dos padrões identificados

são de responsabilidade da gestão da Universidade e dos professores, que podem utilizar as informações apresentadas neste trabalho da maneira mais conveniente nos espaços acadêmicos, caso tenham interesse pelo resultado deste trabalho. De forma semelhante, não faz parte deste trabalho estimular estudantes e professores a seguirem direções apontadas como resultados deste trabalho.

Faz parte do escopo desta pesquisa construir a ferramenta computacional utilizada para a coleta de dados dos alunos e da observação em sala de aula. Da mesma maneira, tratar, validar e analisar o conjunto de dados disponíveis, a fim de tentar fornecer informações relevantes para viabilizar o entendimento da existência de correlação entre os fatores interno e externo e o desempenho acadêmico.

1.3 Objetivos

Geral

Analisar como os fatores internos e externos do estudante estão relacionados com o desempenho acadêmico (aprovação/reprovação) do estudante, no curso de Engenharia da Computação.

Específicos

- Utilizar a técnica de observação e o aplicativo Cookie, desenvolvido para coletar dados dos estudantes;
- Coletar dados da sala de aula e dos alunos durante um semestre;
- Utilizar mineração de dados para encontrar correlações entre os fatores interno e externo do estudante e o desempenho acadêmico;
- Demonstrar a correlação das variáveis.

1.4 Organização do Trabalho

Esta dissertação está estruturada em cinco capítulos, de forma a possibilitar uma melhor compreensão do documento. Os capítulos a seguir estão organizados da seguinte maneira:

No Capítulo 2, é apresentada a **Fundamentação Teórica**, que retrata o cenário do Ensino Superior do Brasil, a retenção e a evasão nos Cursos de Computação e o desempenho acadêmico; a Mineração de Dados, a importância da análise de dados e da tomada de decisões no ambiente educacional, as aplicações e os desafios na educação. Por fim, os trabalhos relacionados, que tem como finalidade trazer estudos, a partir da sua abrangência, que estão mais próximos deste trabalho.

No Capítulo 3, é descrita a **Metodologia** abordando as etapas, os procedimentos, as estratégias efetuadas para a desenvolvimento do trabalho, bem como as atividades envolvidas. Além disso, encontram-se os métodos utilizados para desenvolver o

Cookie, o aplicativo utilizado pelos estudantes e a técnica de observação em sala de aula.

No Capítulo 4, é apresentada a **Coleta e Análise de Dados**. Neste capítulo, é explicada toda a fase da coleta pela observação e pelo uso do *Cookie*. Também é explicada como foram realizadas as etapas para a mineração dos dados coletados.

No Capítulo 5, encontram-se apresentados os **Resultados** alcançados. Neste capítulo, é exposta a análise dos resultados obtidos pela observação em sala de aula, bem como do aplicativo utilizado pelos estudantes, por meio da mineração de correlação.

Finalmente, no Capítulo 6, apresentam-se as **Considerações Finais** obtidas neste trabalho, bem como as pesquisas futuras.

Os documentos de aceite de participação na pesquisa, Termo de Consentimento Livre e Esclarecido - TCLE docente e discente, devidamente aprovados pelo Comitê de Ética, encontram-se disponíveis nos Apêndices A e B.

Capítulo 2

Fundamentação Teórica

Neste capítulo, encontram-se os conceitos essenciais para o desenvolvimento desse trabalho. Na Seção 2.1, é apresentada uma análise geral sobre o Ensino Superior no Brasil, com foco na Retenção, na Evasão e no Desempenho Acadêmico nos cursos da área de Computação. Na Seção 2.2, é feita uma abordagem sobre Mineração de Dados e a sua relevância na educação. Por fim, na Seção 2.4, o capítulo é encerrado com uma abordagem sobre os trabalhos correlatos.

2.1 Ensino Superior Brasileiro

Com o propósito de entender os estudos que discorrem sobre o desempenho acadêmico na educação superior, é considerado fundamental revisar as definições e os dados apresentados por fontes no Ensino Superior Brasileiro, com a finalidade de poder estabelecer uma análise preliminar mais adequada para os problemas enfrentados neste campo.

No Brasil, há uma convicção generalizada de que o Ensino Superior estabelece um elemento primordial para o desenvolvimento político, econômico e social do País. Por outro lado, percebe-se que não há uma percepção muito clara das suas variedades de funções existentes e da complexidade que ele deve desempenhar. Esta percepção pode ser compreendida a partir de sua história.

Autores como Ribeiro e Outros (2016) relatam que antes da década de 70, “o sistema de ensino superior brasileiro contava com poucas Instituições de Ensino Superior - IES, sendo a maioria de pequeno porte e vocacionada à reprodução de conteúdos” [Ribeiro et al. 2016, p.26]. Para este autor, no Século XX o Ensino Superior Brasileiro começou a crescer, surgiram novas instituições e houve a expansão da oferta de vagas em cursos variados. De acordo com Dourado e Outros [Dourado et al. 2004], tal crescimento levou as diretrizes das políticas do MEC a busca por estímulo, expansão e a modernização do ensino de graduação, levando o Ensino Superior a um processo de expansão acelerado, seja em número de cursos e instituições, seja em números de vagas disponíveis.

Essa expansão tem como ponto positivo, prover oportunidade das IES ofertarem cursos distintos para que o indivíduo tenha oportunidade de adentrar no Ensino Superior. Por outro lado, Dourado e Outros [Dourado et al. 2004] expõe uma preocupação sobre a qualidade dos cursos oferecidos pelas IES, pois a oferta acelerada de matrículas em novas vagas poderia afetar o sistema de qualidade no ensino. Conforme estudos feitos por de Paula Franco [de Paula Franco 2008], dados recentes sobre o cenário da educação superior no Brasil têm levado os especialistas a refletirem sobre como lidar com a situação educacional atual.

Uma pesquisa feita pela Revista Exame (2016)¹ mostra evidências sobre o Censo da Educação Superior. Segundo a revista, os dados foram revelados pelo INEP 2015 e indicam que 11% dos alunos que ingressaram em 2010 no Ensino Superior desistiram no primeiro ano, e até 2014, 49% dos estudantes saíram dos cursos que escolheram fazer em 2010, ou seja, quase metade dos alunos que iniciaram não concluíram.

O Censo de Educação Superior (2016) apresenta novos números com relação à oferta de cursos. Ele aponta que “33.501 cursos de graduação foram ofertados em 2.364 instituições de educação superior no Brasil, em 2015” [Brasil 2016, p.3], sendo estes cursos ofertados tanto na rede pública como na rede privada.

De acordo com o Censo (2017), dados recentes indicam que estes números continuam crescendo. A Figura 2.1 apresenta o cenário do crescimento de matrículas no período de 2007 a 2017, no qual a “matrícula na educação superior aumentou 56,4%” [MEC 2017, p.13].

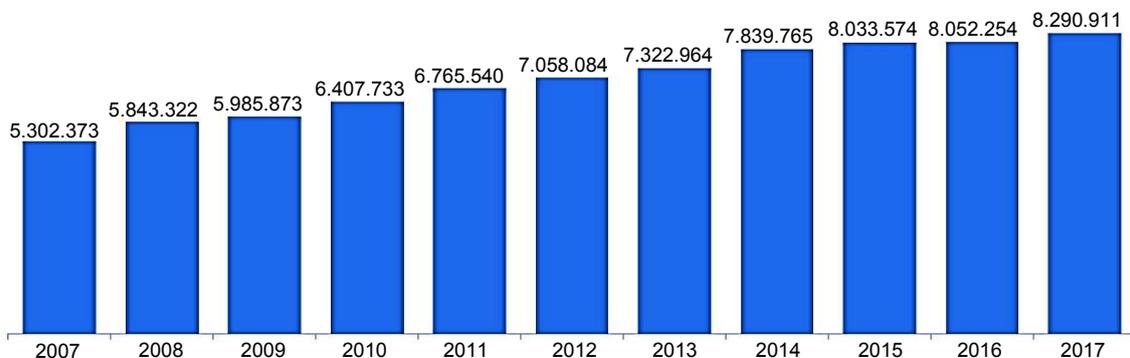


Figura 2.1: Matrículas no Ensino Superior (2007-2017).
Fonte: Adaptada de Censo de Educação Superior [MEC 2017].

Segundo o Censo [MEC 2017], em 2017 o número de matrículas no Ensino Superior atingiu aproximadamente 8,3 milhões (ver Figura 2.1), obtendo uma média anual de crescimento de 4,6%. Por outro lado, face a esse crescimento, encontram-se as dificuldades dos estudantes que ingressaram em um curso de nível superior, concluí-

¹<http://exame.abril.com.br/brasil/10-numeros-que-mostram-como-esta-o-ensino-superior-no-brasil/> Acesso em 24 de Setembro de 2017

rem o mesmo. De acordo com o Censo [MEC 2017], aproximadamente 1,2 milhões de estudantes concluíram o curso de graduação em 2017.

Essas desistências trazem prejuízos para as instituições, aos professores e aos próprios alunos. De acordo com Silva Filho e Outros (2007), “as perdas de estudantes que iniciam mas não terminam seus cursos são desperdícios sociais, acadêmicos e econômicos” [Silva Filho et al. 2007, p.642].

Diante desse cenário, pesquisadores como Sales e Outros [Sales et al. 2014] discutem sobre o processo da desistência entre os estudantes, que tem como consciência a problemática das evasões e/ou retenções, no Ensino Superior.

2.1.1 Retenção e Evasão no Curso de Computação

No país, algumas expressões são utilizadas para qualificar o resultado do corpo discente, seja pelo seu sucesso ou o insucesso. O sucesso está diretamente ligado à sua diplomação, ou seja, a conclusão do curso de graduação. Já o insucesso se refere àqueles estudantes que, ao findar o semestre, não conseguem avançar para o semestre seguinte ou que, por consequência do seu desempenho (aproveitamento) baixo, acabam deixando disciplinas em atraso [Menezes Silva 2010]. Esta situação leva o estudante a prolongar o tempo de conclusão do curso ou até mesmo a desistência, termos muito utilizados como retenção e evasão.

Autores como Manhães [Manhães 2015] explicam que a retenção retrata a situação no qual o estudante, que matriculado em uma instituição, continua cursando além do prazo máximo estabelecido para a sua conclusão curricular. Sales e Outros [Sales et al. 2014] trazem uma definição na qual a retenção está relacionada com a reprovação nas disciplinas, levando o estudante ao atraso para concluir o curso.

Já a evasão, de acordo com Manhães [Manhães 2015], é aquela na qual o estudante, que matriculado em um determinado curso do ano letivo, não faz a matrícula no ano posterior, independente do seu desempenho acadêmico ter sido aprovado ou reprovado. Para Sales e Outros [Sales et al. 2014], a retenção, em muitos casos, fomenta a evasão pois é o ato em que um estudante deixa de fazer sua matrícula regularmente no semestre.

A Comissão Especial de Estudos sobre a Evasão nas Universidades Públicas Brasileiras [MEC 1997] traz em seu documento que a evasão pode ser caracterizada de três maneiras: a) evasão do curso: quando o estudante se desliga do curso superior, mas mantém o vínculo com a instituição; b) evasão da instituição: se refere ao desligamento do estudante da instituição na qual está matriculado e; c) evasão do sistema: nesse caso o estudante abandona o ensino superior.

As três maneiras de evasão citadas anteriormente são preocupantes e merecem atenção, pois segundo Fagundes e Outros [Fagundes et al. 2014], a evasão de modo geral afeta: i) o estudante, levando-o ao atraso do curso e, até mesmo a desistência; ii) a instituição, pois afeta a qualidade e a avaliação da mesma e; iii) o país, visto que aumentam os gastos públicos com a educação.

Diversos são os motivos que podem levar o estudante à evasão e/ou retenção de uma disciplina ou de um curso. O Documento Orientador [MEC 2014] expõe algumas causas que incidem sobre estes fatores, tais como: acesso às instituições, condições e fatores socioeconômicos, os problemas de aprendizagem ou dificuldades nas disciplinas, a repetência ou o desempenho acadêmico insuficiente, as práticas pedagógicas, a exigência dos professores, o excesso de matérias/disciplinas por período do curso, entre outros.

Estes fatores podem incidir em qualquer curso. Contudo, os autores Pascoal e Outros [Pascoal et al. 2016] e Hoed [Hoed 2016] afirmam que as áreas de Computação, Matemática e Ciências estão entre as maiores taxas de evasão, chegando a ficar acima da média nacional. Os autores afirmam ainda que, nos cursos de Computação, as dificuldades apresentadas pelos estudantes encontram-se nas disciplinas voltadas para programação, possuindo altos índices de reprovação, levando possivelmente ao impedimento ou até mesmo a continuidade dos estudantes no curso [Pascoal et al. 2016, Hoed 2016].

No curso de Engenharia da Computação da UEFS não é diferente. Segundo Bittencourt e Outros [Bittencourt et al. 2013], os dados apresentados entre 2009 e 2011 mostram que a taxa de evasão foi de 20%. Já a retenção oscilou entre 50% e 75% nas disciplinas iniciais de programação.

Dentre os motivos que podem causar a retenção e/ou evasão apresentados pelo Documento Orientador [MEC 2014], Hoed (2016) aponta a dificuldade de adaptação dos alunos para desenvolverem o raciocínio lógico e problemas ligados a didática no ensino de algoritmos, como os motivos de maior destaque nos cursos de computação [Hoed 2016].

Os problemas/dificuldades de aprendizagem, por exemplo, assim como as práticas e/ou exigências pedagógicas podem também impactar no desempenho do estudante ao longo do processo de ensino aprendizagem. Dessa forma, é importante e necessário compreender o que se entende por desempenho acadêmico.

Por isso, neste trabalho, o desempenho acadêmico se refere ao resultado (aproveitamento) obtido pelo estudante, sendo ele aprovado ou reprovado, após ter seus conhecimentos adquiridos avaliados por meio de exame (i.e. teste, prova, entre outros), durante seu processo de formação, conforme explicado na Subseção 2.1.2.

2.1.2 Desempenho Acadêmico

O desempenho acadêmico tem sido assunto de discussão na literatura científica, e entendido que, apesar de não ser uma problemática recente, é comum ainda nos dias atuais, pela importância da temática em qualquer nível educacional [de Cássia Martinelli e Genari 2009]. Conhecer os fatores internos e externos e o impacto no desempenho acadêmico têm sido interesse da comunidade acadêmico-científica [Miranda et al. 2015] e [Moleta et al. 2017].

Fatores como a trajetória que o estudante percorre antes de ingressar na faculdade, características socioeconômicas, a compatibilidade do curso com o mercado de trabalho, a idade, o sexo, a determinação, entre outros têm sido estudados pelos autores da Silva [da Silva 2015] e de Leon e Menezes Filho [de Leon e Menezes-Filho 2001] como variáveis que podem impactar o desempenho acadêmico. Outros fatores como equipe de docentes, a estrutura (e.g. biblioteca, sala de aula) da IES, também podem ser determinantes para o desempenho acadêmico [Miranda et al. 2015].

Pesquisas recentes feitas por Porto e Gonçalves [Porto e Gonçalves 2017] destacam que a aprendizagem do estudante está relacionada com a participação nas atividades acadêmicas dentro ou fora da sala de aula. Os autores Wolf-Wendel e Outros [Wolf-Wendel et al. 2009], Moleta e Outros [Moleta et al. 2017] e Fior e Mercuri [Fior e Mercuri 2018] apontam características como a dedicação nos estudos, as atividades elaboradas em grupo bem como a execução de tarefas extracurriculares, a participação em palestras, as interações entre os estudantes, e entre eles e o professor, como fatores que podem afetar o desempenho discente.

Os autores Fior e Outros [Fior et al. 2013] buscam encontrar relação entre o envolvimento acadêmico e as atividades obrigatórias, como fazer anotações durante as aulas, realizar atividades e trabalhos de casa, prestar atenção nas aulas, estudar para avaliações, entre outros, e não obrigatórias como participação de congressos, projetos, organização de atividades artísticas e culturais, de monitoria e/ou tutoria, entre outros, com estudantes das áreas Ciências Humanas e Sociais, Ciências Exatas e Ciências Biológicas. Os autores Costa e Outros [Costa et al. 2014] investigam a relação entre a participação e o desempenho com estudantes de licenciaturas nos cursos de Engenharia Civil, Química, Eletrotécnica, Informática e de Sistemas.

Alguns estudos como os de Campos e Outros [Campos et al. 2012] buscam variáveis, através da utilização das Tecnologias da Informação e Comunicação (TIC), que vêm crescendo cada vez mais, que possam impactar no desempenho do estudante. Estes autores buscam encontrar relações entre as habilidades tecnológicas e o rendimento discente com estudantes dos cursos de Educação a Distância (EaD); Silva e Outros [Silva et al. 2012] e Rangel e Miranda [Rangel e Miranda 2016] buscam verificar a relação entre o uso das Redes Sociais e o desempenho de discentes de graduação do curso de Contabilidade. Já os autores Kirschner e Karpinski [Kirschner e Karpinski 2010] buscam a relação entre o uso da Rede Social *Facebook* e o desempenho acadêmico de estudantes dos cursos de graduação e pós-graduação.

Porém, Gouveia e Outros [Gouveia et al. 2010] alertam que independente do que pode impactar no desempenho acadêmico dos estudantes, ele acaba sendo mensurado por meio de avaliações que vão de zero a dez, que tentam demonstrar o nível de conhecimento e as habilidades adquiridas pelos estudantes no contexto educacional. Sendo assim, a avaliação aplicada ao estudante é o instrumento mais utilizado para medir a eficácia do desempenho em um curso de ensino superior [Nogueira et al. 2013].

Corroborando com Gouveia e Outros [Gouveia et al. 2010], Munhoz [Munhoz 2004]

afirma que “o termo desempenho envolve a dimensão da ação e o rendimento é o resultado da sua avaliação, expresso na forma de notas ou conceitos obtidos pelo sujeito em determinada atividade” [Munhoz 2004, p.37].

Dessa forma, o desempenho está diretamente associado ao resultado e aos fatores que o influencia, tais como determinação, nível de conhecimento, rendimento do estudante nas avaliações, entre outros. Existe ainda a possibilidade de que haja interferência de outros fatores no aluno (i.e localização das universidades, metodologias aplicadas pelo professor em sala, métodos de estudo, entre outros), que não são muito evidentes e que podem estar relacionados com os aspectos do desempenho e, como consequência, com a evasão e com a retenção.

2.2 Mineração de Dados

Dados têm se tornado cada vez mais difíceis de serem analisados por profissionais qualificados. Tal dificuldade é devida a quantidade, ao tipo e as dimensões de análises dos dados, visto que este é um processo que vai além de coleta, processamento e saída de informações. Os dados, nesse caso, precisam ser transformados em conhecimento para apoiar em futuras tomadas de decisão.

Para ajudar nesse processo de análise em grandes bancos de dados, surge a Mineração de Dados (MD) ou do inglês, *Data Mining*. Hand e Outros [Hand et al. 2001] definem a MD como o processo de descoberta de estruturas importantes, inesperadas ou valiosas em grandes conjuntos de dados. Sua principal funcionalidade está na aplicação de vários métodos e algoritmos com intuito de descobrir e extrair padrões de dados armazenados [Sachin e Vijay 2012].

Na literatura, existem vários entendimentos para o termo mineração de dados, alguns dos quais serão abordados. Porém, para este trabalho, em concordância com [Braga 2005], a descoberta do conhecimento é dividida em várias etapas e a mineração de dados, assim como todas as suas etapas são uma instância do KDD (do inglês, *Knowledge Discovery in Databases*).

De acordo com os autores Fayyad e Outros [Fayyad et al. 1996] e Sachin e Vijay [Sachin e Vijay 2012], a mineração de dados é uma etapa no processo do KDD que consiste em aplicar algoritmos de análise e descoberta de dados que produzem uma enumeração particular de padrões (ou modelos) sobre os dados. Esta técnica tem como objetivo encontrar padrões em diversos tipos de dados agrupados, gerando nova fonte de conhecimento com capacidade de identificar relações existentes entre eles. Sendo assim, tem um foco rico devido à sua importância na tomada de decisões.

Os autores Hand [Hand 2007] e Witten e Outros [Witten et al. 2016], afirmam que a mineração de dados é definida como o processo de descoberta de padrões, estruturas, inesperadas ou valiosas em grandes conjuntos de dados.

Segundo Sachin e Vijay [Sachin e Vijay 2012], a mineração de dados e a descoberta de conhecimento em bancos de dados são tratadas como sinônimos. No entanto,

no livro de Braga [Braga 2005, p.15], *Introdução a Mineração de Dados*, o autor assegura que “a mineração de dados está inserida em um processo maior denominado descoberta de conhecimento em banco de dados (KDD)”.

Na opinião de Tavares e Outros [Tavares et al. 2016], o KDD aborda várias fases de investigação e depuração com a finalidade de encontrar padrões em informações úteis a partir de grandes bases de dados. O processo que é dividido em várias etapas pode ser visto na Figura 2.2.

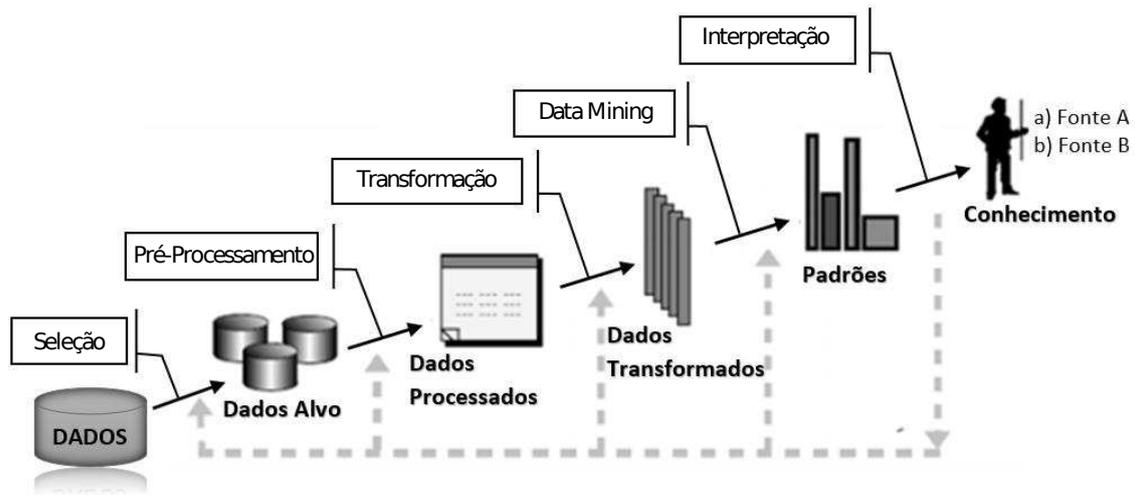


Figura 2.2: Etapas da extração de conhecimentos de dados.
Fonte: Adaptada de Sachin e Vijay [Sachin e Vijay 2012].

O KDD, como pode ser observado na Figura 2.2, é baseado em um ciclo variável, no qual cada fase pode ser revisitada várias vezes até que o objetivo final seja alcançado. Segundo Han e Outros [Han et al. 2011]; Pasta e Outros [Pasta 2011], este processo de conhecimento é uma sequência iterativa de passos, as quais estão explicadas a seguir:

1. **Seleção de Dados:** define o domínio que será executado o processo de descoberta, seleção e coleta de dados. Nesta fase, segundo Han e Outros [Han et al. 2011], os dados relevantes para a tarefa de análise são recuperados do banco de dados.
2. **Pré-processamento:** é responsável pelo tratamento dos dados de entrada, cuja finalidade é analisar, integrar, transformar e limpar os dados, visando a melhoria da qualidade dos dados. Para isso, segundo [Fayyad et al. 1996], explica que estas operações são necessárias para modelar ou explicar o ruído, decidindo sobre estratégias para lidar com campos de dados ausentes e contabilizando informações de sequência de tempo e mudanças conhecidas.

3. **Transformação de Dados:** nesta etapa os dados são transformados e consolidados em formulários apropriados para mineração [Han et al. 2011]. Com redução de dimensionalidade ou métodos de transformação, o número de variáveis em consideração pode ser reduzido, ou representações invariantes para os dados podem ser encontradas, sem sacrificar sua integridade [Fayyad et al. 1996].
4. **Mineração de Dados:** tem como objetivo definir qual aplicação de algoritmos serão utilizados na extração de padrões. Segundo Fayyad e Outros [Fayyad et al. 1996], esta fase busca por padrões de interesse em uma forma representacional particular ou um conjunto de tais representações, incluindo regras de classificação ou árvores, regressão e agrupamento.
5. **Interpretação/Conhecimento:** é a descoberta do conhecimento de maneira que o usuário possa entender e interpretar os resultados obtidos. Para Han e Outros [Han et al. 2011], nesta fase são utilizadas técnicas de visualização e representação das informações a fim de apresentar o conhecimento explorado aos usuários.

Esta técnica abre espaço ainda para ser utilizada dentre outras áreas, na Educação. Segundo Costa e Outros [Costa et al. 2013] estudiosos têm apresentado interesse em utilizar a mineração de dados para compreender os dados em ambiente educacional, originados de estudantes e professores. Diante desse contexto, surgiu a área de pesquisa conhecida como Mineração de Dados Educacionais do inglês, *Educational Data Mining* (EDM) [Baker et al. 2011].

2.3 Mineração de Dados Educacionais

Com o aparecimento dos cursos a distância e da utilização cada vez mais frequentes dos meios computacionais, seja para a aprendizagem, seja para o gerenciamento, pesquisadores têm mostrado interesse em investigar dados educacionais a fim de procurar melhorias para o ensino [Baker et al. 2011].

Estudos feitos por Backer e Outros [Baker et al. 2011] abordam que, principalmente estudiosos de Inteligência Artificial Aplicada à Educação têm utilizado técnicas de mineração de dados para tentar responder, por exemplo, quais os fatores que podem afetar a aprendizagem do estudante? Diante desse cenário, pesquisas voltadas para essa área vêm sendo feitas, a fim de compreender quais variáveis podem ter relação com o desempenho do estudante.

Desse modo, os autores Backer e Outros [Baker et al. 2010] e Algarni [Algarni 2016] definem a EDM como a área de investigação científica voltada para o desenvolvimento de métodos a fim de criar descobertas nos dados que surgem dos meios educacionais para entender melhor os alunos e os ambientes em que eles aprendem.

De acordo com Sachin e Vijay [Sachin e Vijay 2012], a aplicação da mineração de

dados no sistema educacional é um ciclo interativo de formação, teste e refinamento de hipóteses, conforme apresentado na Figura 2.3.

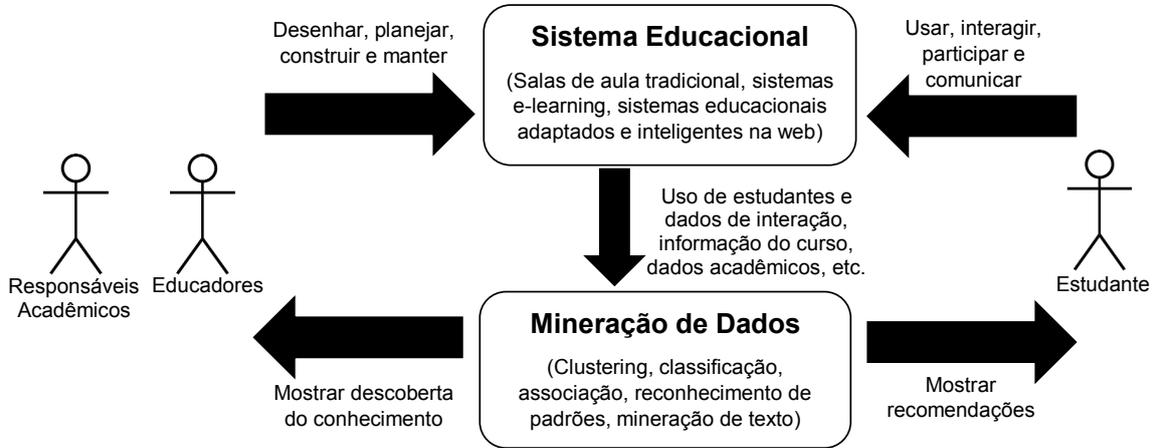


Figura 2.3: Ciclo de aplicação de mineração de dados em sistemas educacionais.
Fonte: Adaptada de Sachin e Vijay [Sachin e Vijay 2012].

Como pode ser observado na Figura 2.3, o sistema educacional é alimentado pelos dados das pessoas (e.g. estudantes, professores, gestores) que interagem com o ambiente através de planejamentos, participações, entre outros, no intuito de facilitar, orientar e melhorar o ensino-aprendizagem. Enquanto a EDM utiliza os dados coletados no meio para transformá-los em conhecimento e auxiliar em tomadas de decisão.

Com isso, a EDM procura adaptar as tarefas já existentes na mineração de dados no ambiente educacional disponíveis por professores, estudantes, plataformas de interação, como Ambiente Virtual de Aprendizagem (AVA), Sistema Tutor Inteligente (SIT), de maneira que os dados sejam melhor compreendidos e analisados [Costa et al. 2013].

2.3.1 Tarefas da Mineração de Dados Educacionais

Segundo Castro e Ferrasi [Castro e Ferrari 2016] a mineração de dados possui funções que são utilizadas para especificar os tipos de informação que serão obtidas nas tarefas de mineração. Na mineração de dados educacionais não é diferente, pois as tarefas são originalmente da área de mineração de dados [Witten et al. 2016].

A seguir estão apresentadas uma sistematização das principais sub-áreas da pesquisa em mineração de dados educacionais.

- **Predição:** também conhecido como Aprendizado Supervisionado, tem como objetivo fazer previsões. A predição faz inferência nos dados a fim de prever o valor de um atributo através de outros de um subconjunto da base de dados [Pasta 2011] e [Costa et al. 2013].

De acordo com [Costa et al. 2013], as vantagens em utilizar a predição em EDM estão principalmente em poder estudar quais os fatores principais para a predição e poder apoiar no desenvolvimento e aplicação de atividades.

- **Classificação:** para Lu e Outros [Lu et al. 1995]; Costa e Outros [Costa et al. 2013], a classificação é o processo de encontrar as propriedades comuns entre diferentes entidades e classificá-las em classes. Assim, os resultados obtidos são frequentemente expressos sob a forma de regras, denominadas de regras de classificação. O objetivo da classificação é prever com precisão a classe de destino para cada caso nos dados [Krishnaiah et al. 2014].

Os modelos de classificação são testados comparando os valores previstos aos valores alvo conhecidos em um conjunto de dados de teste. Os dados históricos de um projeto de classificação são divididos em dois conjuntos de dados: um para construir o modelo; o outro para testar o modelo [Krishnaiah et al. 2014].

Nesta tarefa, os algoritmos mais utilizados são: i) árvores de decisão: utilizam treinamento supervisionado para classificação e predição dos dados e; ii) máquina de vetores de suporte, utiliza um hiperplano (conjunto de treinamento) para separar classes [Costa et al. 2013].

- **Regressão:** segundo Dias [Dias 2002] a regressão é utilizada para identificar um valor para uma variável contínua desconhecida. Com isso, o valor de uma determinada variável pode ser estimado fazendo análise das outras variáveis. Nesta tarefa os algoritmos mais utilizados são regressão linear, redes neurais e máquinas de vetores de suporte para regressão [Costa et al. 2013].

Os autores Camilo e Silva [Camilo e Silva 2009], trazem um exemplo no qual pode ser estimada a pressão ideal de um paciente tomando por base a idade, o sexo e a massa corporal. Após ter analisado os dados, o modelo é capaz de dizer qual será a pressão ideal para um novo paciente. Em Rangel e Miranda [Rangel e Miranda 2016], é utilizada regressão linear para encontrar relação entre o uso de redes sociais e o desempenho acadêmico.

- **Clusterização:** são técnicas de aprendizado Não-Supervisionado tendo como objetivo encontrar grupos ou objetos que possuem similaridades [Baker et al. 2011] e [Costa et al. 2013]. Diferente da classificação que prediz o valor de uma variável, o *cluster* identifica grupos de dados homogêneos ou heterogêneos [Camilo e Silva 2009].

Para Castro e Ferrari [Castro e Ferrari 2016] um *cluster* pode ser definido como uma coleção de atributos similares uns aos outros e dissimilares aos atributos pertencentes a outros *clusters*. Por exemplo, encontrar grupos de estudantes

a fim de investigar suas semelhanças e diferenças ou grupos por padrões de comportamento.

De acordo com Costa e Outros [Costa et al. 2013] algoritmos genéticos, o K-means ou do tipo Fuzzy são muito utilizados nessa tarefa. Os autores Pimentel e Outros [Pimentel et al. 2003] utilizaram técnicas de clusterização com o algoritmos K-means para identificar grupos similares de aprendizes.

- **Mineração de relações:** a finalidade é descobrir, em um banco de dados com muitas variáveis, relações entre elas. Para isto, é feita uma investigação para buscar quais variáveis estão fortemente relacionadas, ou seja, utiliza análise estatística para identificar a força e a relação entre variáveis [Costa et al. 2013] e [Amaral 2016].

A mineração de relações está dividida em quatro tipos. A regra de associação é uma técnica que segundo [Costa et al. 2013] é distribuída em outras variações: i) correlações; ii) Padrões Sequenciais e; iii) causas, as quais serão descritas a seguir.

- **Mineração de Regras de Associação:** é a descoberta de regras e conexões entre os dados, ou seja, identifica a relação entre atributos e gera resultados esperados.

De acordo com Witten e Outros [Witten et al. 2016], as regras de associação podem prever qualquer atributo, não restrito somente a classe, dando a possibilidade de prever combinações de atributos também. Por exemplo, em EDM a regra de associação pode ser utilizada em um banco de dados que contém notas de estudantes. Nesse ambiente seria possível derivar regras que definem uma porcentagem “X” de estudantes com bom desempenho em uma determinada disciplina [Costa et al. 2013].

Existem muitos algoritmos que são utilizados nessa tarefa tais como, regras de associação, padrões sequenciais, porém o mais utilizado, de acordo com [Costa et al. 2013] é o Apriori. Este algoritmo combina a busca com estrutura de árvore, o que o torna eficiente.

- **Mineração de Correlações:** essa técnica tem como objetivo encontrar correlações lineares, seja ela positiva ou negativa, entre as variáveis [Baker et al. 2011]. Por exemplo, através de dados coletados, analisar como os fatores internos e externos estão relacionados com o desempenho acadêmico. Os coeficientes de correlação de *Pearson* e de *Spearman* são muito utilizados na mineração de dados educacionais [Dancey e Reidy 2018].

Em Campos e Outros [Campos et al. 2012] é utilizado o coeficiente de correlação de *Pearson* para buscar relações entre habilidades tecnológicas em alunos de cursos EAD e o desempenho acadêmico. As autoras [de Cássia Martinelli e Genari 2009] utilizaram o coeficiente de *Spearman*

para investigar as relações entre o desempenho de estudantes do ensino fundamental e as suas orientações motivacionais.

- ***Mineração de Padrões Sequenciais:*** o objetivo principal é encontrar associação temporal entre determinados eventos e o impacto destes eventos no valor de um atributo [Baker et al. 2011].
- ***Mineração de Causas:*** nesse caso, os algoritmos desenvolvidos são utilizados para investigar se um evento causa outro evento, para isso são utilizados padrões de covariância [Baker et al. 2011].
- **Destilação de dados para facilitar decisões humanas:** tem como finalidade encontrar uma maneira de permitir que os pesquisadores identifiquem os classifiquem os recursos nos dados com facilidade. É uma área de interesse de EDM, uma vez que é utilizada para identificar comportamento ou colaboração dos estudantes [Costa et al. 2013] e [Algarni 2016].

Os autores Backer e Outros [Baker et al. 2011] explicam que os métodos utilizados nesta sub-área da mineração de dados educacionais facilitam a maneira como são visualizadas as informações que estão inseridas nos dados coletados, por exemplo, por sistemas educacionais. Um tipo de visualização muito utilizada é curva de aprendizagem, pois apresenta o nível de aprendizagem de um estudante.

- **Descoberta de Modelos:** desenvolve um modelo de fenômeno usando *clustering*, previsão ou engenharia do conhecimento [Algarni 2016]. Em seguida, esse modelo é utilizado como ponto de partida em outra análise. Por exemplo, a descoberta de relações entre comportamentos dos estudantes ou variáveis contextuais e análise de questões de pesquisas em vários contextos [Costa et al. 2013] e [Algarni 2016].

As tarefas de mineração de dados podem ser aplicadas por meio de técnicas distintas e ao escolher uma técnica, é importante ter conhecimento sobre ela, a fim de obter ajuda no estudo dos problemas apresentados [Dias 2002].

Nesta dissertação, são apresentados alguns projetos publicados por autores como os de Costa e Outros [Costa et al. 2014] Silva e Outros [Silva et al. 2012], Campos e Outros [Campos et al. 2012], Rangel e Miranda [Rangel e Miranda 2016], Kirschner e Karpinski [Kirschner e Karpinski 2010] e Rui Wang e Outros [Wang et al. 2014]. Estes trabalhos têm utilizado estratégias como, questionário e aplicativo instalado em *Smartphone*, para coletar dados e técnicas de mineração de dados para buscar nos dados, correlações entre os fatores internos e externos à sala de aula e o desempenho acadêmico do estudante. Estas pesquisas encontram-se citadas nos Trabalhos Relacionados, Seção 2.4.

2.4 Trabalhos Relacionados

Os autores Costa e Outros [Costa et al. 2014] utilizaram o Inventário de Envolvimento Acadêmico de Estudantes do Ensino Superior para coletar dados de 360 estudantes do 1º ano das licenciaturas de Engenharia Informática, Civil, Química, Eletrotécnica e de Sistemas. O intuito é encontrar relação entre as medidas de envolvimento e o fato de os estudantes já terem pensado em desistir do curso que estuda. Os resultados apontaram para uma diferença estatística entre o envolvimento emocional, o que significa dizer que os estudantes com menor envolvimento afetivo à escola e ao curso, tem uma maior tendência em desistir. Com a relação as outras dimensões, não foram encontrados resultados significativos

Foram analisados também se há uma relação entre o envolvimento e o desempenho acadêmico. Os resultados apresentaram uma relação positiva e baixa entre as três dimensões, i) Envolvimento Comportamental, ii) Envolvimento Emocional e; iii) Envolvimento Cognitivo, porém o coeficiente considerado correlação moderada foi o envolvimento comportamental. Para os autores, este é um resultado que envolve a participação do estudante e o rendimento do mesmo nas avaliações.

Os autores analisaram ainda o impacto das três dimensões de envolvimento e o desempenho acadêmico, utilizando regressão linear. Os resultados apresentaram um valor estatisticamente significativo, revelando a existência de impacto positivo entre o envolvimento comportamental e o desempenho, as demais dimensões não apresentaram impactos significativos. Os autores concluem que o estudo feito traduz a relevância do envolvimento dos estudantes com atividades curriculares e extracurriculares e a aprendizagem. Já o envolvimento comportamental é relacionado com as horas dedicadas aos estudos, o esforço na participação das atividades dentro ou fora da sala de aula. Para trabalhos futuros, os autores recomendam utilizar uma amostra mais diversificada de estudantes de Engenharia de outras instituições ou ainda de períodos diferentes.

No trabalho de Silva e Outros [Silva et al. 2012], os autores utilizaram um questionário que coletou dados de 105 estudantes do curso de Graduação em Ciências Contábeis da Universidade Federal de Alagoas (UFAL). O objetivo foi de investigar se há relação entre o uso de redes sociais e o desempenho acadêmico. A coleta de dados foi feita no período de agosto a novembro de 2011.

Os autores explicam que foram realizadas análise qualitativa e quantitativa, as quais utilizou-se métodos de estatísticas descritivas e análise de variância multivariada (MANOVA). Os autores elaboraram três hipóteses a fim de confirmar ou refutar a hipótese de que há relação entre o uso de redes sociais e o desempenho, são elas: i) o uso de redes sociais pode ser relacionado ao desempenho acadêmico de estudantes de graduação em contabilidade; ii) o uso de redes sociais tem impacto diferente no desempenho acadêmico de estudantes de graduação de contabilidade calouros e veteranos e; iii) o uso de redes sociais impacta o desempenho acadêmico de alunos veteranos.

Segundo os autores, os resultados obtidos apontam que não é possível afirmar que o uso das redes sociais impacta no desempenho do estudante. A análise qualitativa não pode ser realizada, pois não havia número suficiente de respondentes para realizar o teste de hipótese. Na conclusão, os autores se preocupam com a carência dessa temática, uma vez que existe a necessidade de mais estudos que possam verificar a relação entre o uso das redes sociais e o rendimento discente. Além disso, eles sugerem para novas pesquisas, fazer um estudo longitudinal, verificando a evolução do estudante ao longo do tempo e monitoramento do uso das redes sociais e, comparar o uso das redes entre calouros, veteranos, pessoas mais velhas e mais novas, como também estender a pesquisa para outros cursos e instituições privadas.

Os autores Campos e Outros [Campos et al. 2012] buscaram verificar se existe relação entre as habilidades tecnológicas dos estudantes e o desempenho nos cursos de EaD. Para isso, foi feita uma pesquisa com 1105 estudantes do curso de especialização à Distância, na UFC Virtual, no qual mantém suas atividades realizadas dentro do ambiente virtual Moodle.

Para ajudar na coleta de dados, foi aplicado um questionário *online*, no qual foi composto por 16 questões e dividido em 3 três blocos que trataram sobre o perfil demográfico, tecnológico e as habilidades tecnológicas dos estudantes. O desempenho dos estudantes, ficou definido pelos autores, como a média das três notas nas disciplinas (Sala Ambiente de Introdução ao Curso e ao Ambiente Virtual, Sala Ambiente Oficinas Tecnológicas e Sala Ambiente Tópico Especiais - A Qualidade da Educação Escolar). Para realizar as análises estatísticas descritiva, utilizou-se o sistemas SPSS, bem como foi utilizada a tabulação cruzada com qui-quadrado de *Pearson*, análise de confiabilidade e fatorial.

Nos resultados, os autores explicam que o estudo apresentou uma correlação baixa, porém significativa, entre as habilidades básicas (i.e. navegação na Internet; utilização de *e-mail*; planilha eletrônica, editor de textos entre outros) e o desempenho do estudante. E, foram encontrados resultados não significativos estatisticamente para o uso das habilidades virtuais (i.e. participação de fóruns de discussão; salas de bate-papo (*chats*); *blogs*; e participação em comunidades virtuais) e o desempenho acadêmico. Para trabalhos futuros, os autores recomendam a utilização de outras variáveis que caracterize o desempenho acadêmico, que não seja a nota. Além disso, eles sugerem utilizar outros aspectos que possam ser caracterizados como habilidades tecnológicas.

Rui Wang e Outros [Wang et al. 2014] apresentam um projeto no qual foi desenvolvido um aplicativo *Android* para ser utilizado no *smartphone* dos estudantes de Dartmouth College - Nova Inglaterra. O aplicativo serve como sensor para coletar dados diariamente do aluno e avaliar o impacto da carga de trabalho no estresse, no sono, nas atividades diárias, no humor, na sociabilidade, no bem-estar mental e no desempenho acadêmico. A pesquisa envolveu um total de 48 estudantes, na disciplina de Programação para *Android*, nos cursos de graduação e pós-graduação, em um período de 10 semanas.

A aplicação oferece dois tipos de funcionalidades: 1) dados coletados automaticamente, como duração de conversas, sono, atividades físicas, localização por meio de GPS e *Bluetooth* e; 2) dados introduzidos manualmente pelos estudantes através dos componentes *MobileEMA* e *SurveyMonkey*². O EMA (*Ecological Momentary Assessment*) faz a coleta de dados, através de questionários que servem para investigar dados relacionados ao estresse e humor, e o *SurveyMonkey* é utilizado para administrar pesquisas sobre a saúde e o comportamento. Além disso, os autores tiveram acesso aos dados acadêmicos dos estudantes, a fim de ajudar na avaliação do desempenho acadêmico. Os dados coletados na pesquisa são enviados para uma nuvem toda vez que o celular tiver carregado e conectar a uma rede Wi-Fi.

Para detectar o sono, a atividade física e as conversas, os autores utilizaram classificadores. Com isso, foram encontradas a duração da conversação e a frequência durante o dia, no qual apresentam uma correlação positiva com as notas dos estudantes. Também foram encontradas ligações entre essas medidas (conversas e mobilidades) e a média das notas cumulativas, o que implica dizer que os dados coletados pelos sensores durante o período da pesquisa estão associados ao desempenho geral da faculdade.

Segundo os autores, os resultados do *StudentLife* mostram uma série de correlações significativas entre os dados coletados do sensor de *smartphones*, a saúde mental e o desempenho acadêmico do corpo estudantil. Eles relatam que a conversação e a mobilidade interna dos alunos têm forte relação com o desempenho acadêmico. Os autores afirmam que a medida em que o semestre avança e a carga de trabalho aumenta, o estresse aumenta, enquanto o sono, as conversas e o bem estar mental diminuem. Na conclusão, os autores questionam se é possível encontrar tendências em um corpo estudantil diferente, utilizando o *smartphone* como sensor.

A comparação existente entre os trabalhos relacionados consiste em encontrar relações entre as variáveis coletadas tais como, a participação do estudante, o uso das redes sociais, habilidades tecnológicas e o desempenho acadêmico, coletados através da aplicação de questionários para os estudantes de cursos distintos como, os das Engenharias (Civil, Química, Informática, Sistemas e Eletrotécnica) Ciências Contábeis e especialização a distância EAD.

Já a relação que os trabalhos apresentados nesta pesquisa possuem com esta dissertação consiste na coleta de variáveis (i.e. participação, redes sociais) e no desenvolvimento do aplicativo para coletar dados dos estudantes automaticamente, como apresentado no trabalho de [Wang et al. 2014]. Além disso, esta pesquisa utiliza a análise descritiva e técnicas de mineração de correlação para buscar relação entre o uso de Redes Sociais, a participação do estudante nas aulas e o desempenho acadêmico. Porém, diferente dos trabalhos relacionados, este trabalho utilizou a técnica de observação em sala de aula, para coletar variáveis internas como, a participação do estudante nas aulas e na utilização do aplicativo *Cookie* a fim de coletar variáveis externas, como por exemplo o nome do aplicativo usado no momento da coleta.

²<https://www.surveymonkey.com/>

Capítulo 3

Metodologia

Este capítulo descreve em detalhes os procedimentos metodológicos que são utilizados nesse trabalho. Para a classificação desta pesquisa, optou-se pela estruturação apresentada por Gil [Gil 2002], que determina que os critérios são baseados nos objetivos do trabalho. Sendo assim, o tipo da pesquisa foi considerado a partir dos aspectos relacionados ao nível de profundidade do estudo (objetos), aos procedimentos utilizados para a coleta de dados (técnicas) e à sua abordagem.

Os esquemas utilizados para a coleta de dados foram executadas de duas maneiras: a observação (fontes primárias e secundárias) e por meio do aplicativo instalado no celular dos estudantes, denominado *Cookie*, os quais serão explicadas nas próximas seções deste capítulo.

A observação, segundo Diehl e Tatim [Diehl e Tatim 2004], ajuda a identificar e coletar provas que orientam no comportamento dos indivíduos observados. Nesse caso, são usadas *fontes primárias*, informações obtidas por meio das pessoas envolvidas no projeto, através dos instrumentos de coleta questionário e observação; e *fontes secundárias*, que são provenientes de outras fontes de dados que não são criadas pelo pesquisador, tais como arquivos e banco de dados.

A respeito da abordagem, a pesquisa é de caráter quantitativo, uma vez que pretende-se analisar a relação entre os fatores internos e externos à sala de aula e o desempenho acadêmico de estudantes do curso de Engenharia da Computação da UEFS. A pesquisa quantitativa busca encontrar “a relação entre as variáveis para garantir a precisão dos resultados, evitando contradições no processo de análise e interpretação” [Prodanov e de Freitas 2013, p. 70].

A seguir, serão apresentadas as etapas necessárias para elaboração do projeto. Busca-se elucidar as etapas metodológicas, assim como os esquemas utilizados para a coleta de dados e o desenvolvimento da pesquisa. Estas etapas são necessárias para a execução desta pesquisa, pois inclui a abordagem utilizada para a análise e a apresentação dos resultados. O esquema desse planejamento encontra-se na Figura 3.1.

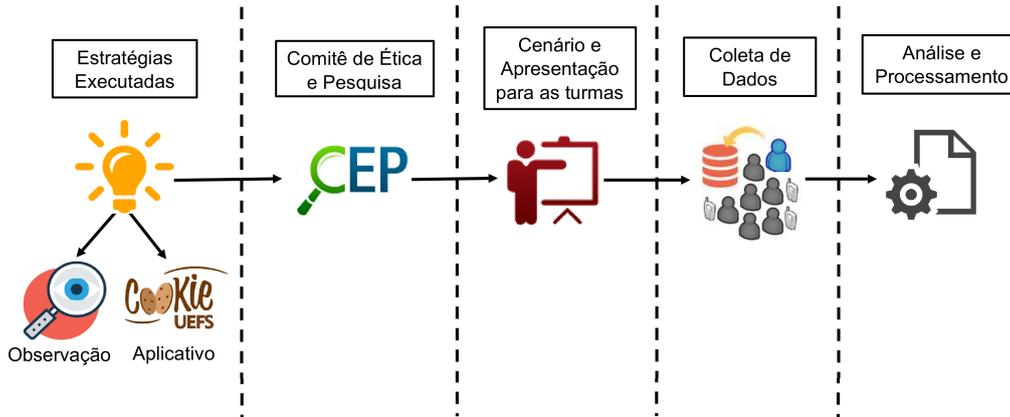


Figura 3.1: Esquema para execução do projeto.
 Fonte: Próprio Autor

Como pode ser visto na Figura 3.1, inicialmente foram pensados e planejados os procedimentos que seriam utilizados na observação por esta pesquisadora, e também no desenvolvimento do aplicativo *Cookie*. Em seguida, o projeto foi submetido ao Comitê de Ética, como meio de garantir a segurança aos participantes, aos dados coletados e à pesquisa.

Uma vez que o projeto foi aprovado pelo Comitê de Ética e as estratégias para a coleta de dados ficaram prontas, o projeto foi apresentado para as turmas e foi dado o início da coleta. Por fim, foi feita a análise dos dados. Todas essas etapas estão explicadas nas próximas seções.

3.1 Observação

A Observação foi realizada pela pesquisadora em todas as aulas, nas turmas citadas na Seção 3.4, do curso de Engenharia da Computação. Para isso, foram registrados dados sobre as aulas, como: participação dos estudantes, horário de chegada e saída, local da sala onde o estudante sentou, entre outros, como pode ser visto na Tabela 3.1.

Tabela 3.1: Variáveis coletadas através da Observação

Dados sobre o Estudante	Dados sobre a Aula
Hora Chegada	Hora Início da Aula
Hora Saída	Hora Final da Aula
Hora Chegada Intermediária	Assuntos
Hora Saída Intermediária	Recursos
Local que Sentou	
Número de Participação	

Fonte: Próprio Autor.

Conforme apresentado na Tabela 3.1, a coleta de dados sobre o estudante se refere aos seguintes itens: a) Hora de Chegada - identifica o horário de chegada do aluno na sala; b) Hora de Saída - identifica o horário de saída do aluno na sala; c) Hora de Chegada Intermediária - identifica o horário que o aluno voltou durante a aula; d) Hora de Saída Intermediária - identifica o horário que o aluno saiu durante a aula; e) Participação é o número de participações do aluno na aula e; f) Local que Sentou é a posição que o aluno sentou na sala (i.e. frente, meio, fundo).

Já na aula, os itens da coleta são: a) Hora de início da aula - identifica a hora de início da aula; b) Hora final da aula - identifica a hora que a aula terminou; c) Assunto - diz respeito sobre o nome do assunto dado pelo professor na aula e; Recursos utilizados - são os recursos usados pelo professor na aula, como por exemplo, exercícios, quadro, projetor, entre outros.

A coleta dos dados, mencionada na Tabela 3.1, foi feita através de uma planilha eletrônica criada para este fim, através dos quais os dados foram anotados conforme os estudantes chegavam, saíam, sentavam e participavam das aulas. Para ajudar no conhecimento dos estudantes, os professores das disciplinas disponibilizaram a lista de presença das turmas.

Além da planilha, foi criado também um mapa de assentos para ajudar na identificação da posição na qual o estudante sentava na sala. O mapa foi utilizado durante todo o período da observação. A Figura 3.2 apresenta o *layout* de assentos da sala de aula.

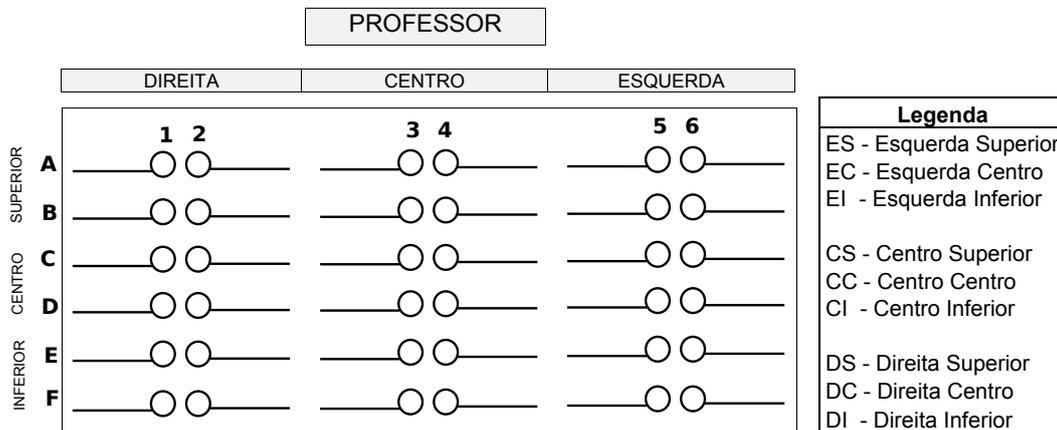


Figura 3.2: Mapa de assentos.

Fonte: Próprio Autor

Como pode ser observado na Figura 3.2, o mapa de assentos ajudou a identificar o local em que o estudante sentou na sala de aula. De forma sistemática, o mapa possui, no centro superior, o professor como referência para orientação dos lados esquerdo, centro e direito da sala.

Sendo assim, foram desenvolvidos três tipos de orientações, conforme encontram-se dispostos na legenda. O encontro da coluna com a linha representa um endereço específico da sala, no qual os números representam os lados direito, centro e esquerdo e as letras representam a frente (superior), o meio (centro) e o fundo (inferior).

Desse modo, um local é definido por um assento, isto é, um estudante ao sentar na posição A1, A2 ou B1, B2 quer dizer que ele está sentado do lado Direito Superior da sala e assim por diante. O processo foi o mesmo para as salas cujo o tamanho e a disponibilidade de cadeiras eram menores.

Conforme o aluno chegava e sentava, o nome dele era anotado no endereço específico, seguindo as orientações do mapa. Estas representações serviram de guia em todo processo de localização do estudante na sala de aula. Além disso, o mapa de assentos ajudou na identificação do estudante que participava/interagia na aula tirando dúvidas com o professor.

3.2 Aplicativo Cookie

O *Cookie* é um aplicativo que foi desenvolvido para plataforma *Android* e instalado no *smartphone* de cada estudante que manifestou o desejo em participar da pesquisa. O aplicativo faz a coleta automática dos dados, ou seja, não é necessário a intervenção do usuário. As variáveis coletadas foram: latitude, longitude, luminosidade do celular, nome do aplicativo em uso pelo estudante, parado/movimento e data da coleta, como mostra a Tabela 3.2.

Tabela 3.2: Variáveis coletadas através do *Smartphone*

Variáveis	Descrição
Latitude	Coordenada Geográfica
Longitude	Coordenada Geográfica
Luminosidade do Celular	Grau de luminosidade do celular
Nome do app em uso pelo estudante	Nome do aplicativo que o aluno usou (i.e. Google)
Parado/Movimento	Identifica se o celular está parado ou em movimento (i.e. 1:Sim 0:Não)
Data da Coleta	Data de registro da coleta

Fonte: Próprio Autor.

Quanto à variável nome do aplicativo, como visto na Tabela 3.2, é registrado apenas o nome do aplicativo usado, portanto, informações sobre o conteúdo, como senhas e as ações do participante no aplicativo não são registradas. Por exemplo, caso o usuário esteja acessando uma página do Banco do Brasil no *browser* do *Chrome* ou *Internet Explorer*, o aplicativo só vai pegar a informação do *browser* (*Chrome*) que está sendo utilizado pelo aluno, ignorando o conteúdo ou a página que está sendo exibida pelo mesmo.

Para desenvolver o *Cookie*, foi necessário levantar os recursos tecnológicos que seriam utilizados. A escolha das tecnologias foi baseada em ferramentas de código aberto pela facilidade de uso e gerenciamento. São elas:

- *NetBeans IDE* - ambiente para o desenvolvimento de aplicações *desktop*, móveis e *web*;
- *Android Studio* - ambiente para o desenvolvimento de aplicativos para dispositivos *Android*;
- *Java* - linguagem orientada a objetos;
- *MySQL WorkBench* - gerenciador de banco de dados;
- *Astah Student* - modelagem do sistema.

O *Cookie* foi desenvolvido para uso apenas em dispositivos com sistema operacional *Android* e tem como objetivo coletar dados dos estudantes, dentro e fora da sala de aula. O nome do aplicativo surgiu da ideia do *cookie de internet*, que são arquivos caracterizados por armazenar, temporariamente, páginas visitadas na rede.

O *Cookie* foi projetado para ser executado automaticamente a partir da sua instalação, não obrigando a interação do estudante com o aplicativo. Atualmente, o *Cookie* encontra-se disponível para baixar na página do ADaM - *Advanced Data Management Reserch Group* ¹.

O aplicativo foi desenvolvido pela autora deste trabalho e com a colaboração de dois estudantes do curso Engenharia da Computação da UEFS, conforme mencionados no TCLE (Discente), Apêndice A. Para a etapa de desenvolvimento, o *Cookie* passou por algumas fases iniciais, tais como:

- **Modelagem**² apresenta o modelo físico e as características do funcionamento e do comportamento do *Cookie*. É através do modelo de dados que o banco de dados é criado, nele estão apresentadas as tabelas, os atributos e os seus relacionamentos;
- **Arquitetura** apresenta como estão organizados os elementos (*Smartphones*, *Web Server*, *Internet*) que compõe o aplicativo.

Dessa maneira, o *Cookie* apresenta, em sua arquitetura, de um lado o Servidor *Web* e, do outro lado, os *smartphones*, os quais se comunicam através de uma rede de internet, como mostra a Figura 3.3.

¹<http://sites.ecomp.uefs.br/adam/projects/cookie>

²<http://bit.ly/2YG022h>

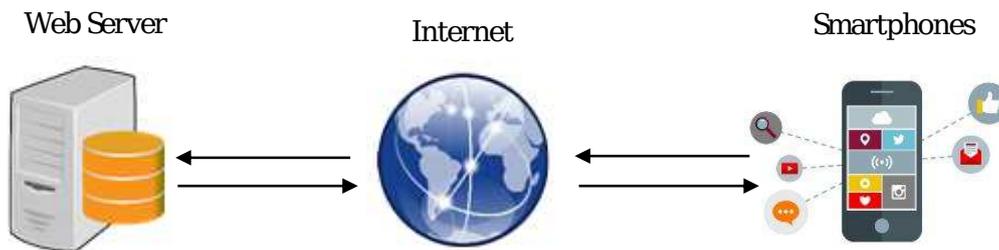


Figura 3.3: Arquitetura do *Cookie*.
Fonte: Próprio Autor

Conforme a Figura 3.3, o *Cookie* tem a capacidade de reunir informações dos alunos através de sensores dos próprios *smartphones* e enviar os dados para um servidor que se encontra localizado na rede.

- **Servidor Web** tem como finalidade receber e armazenar, em um banco de dados, os dados coletados dos *Smartphones* dos estudantes;
- **Aplicativo Cookie** ferramenta instalada no celular dos estudantes com o objetivo de fazer a coleta automática de dados dos alunos a cada minuto e enviá-los para o servidor, quando o celular estiver conectado a uma rede *wi-fi*.

Ao finalizar o desenvolvimento do aplicativo, o *Cookie* passou pela etapa de testes. O teste faz parte do desenvolvimento do sistema e tem como objetivo encontrar falhas no *software* que possam ser corrigidas até que sejam atingidas as necessidades desejadas para utilização.

Foram realizados testes funcionais pelo desenvolvedor e a equipe, e também com usuários específicos convidados para os testes. Estes testes foram feitos antes do uso com os usuários reais, e com os estudantes ao longo do uso durante a pesquisa, à medida em que novos erros foram detectados.

Nesta etapa, o aplicativo foi instalado em *smartphones* com arquiteturas distintas e o sistema operacional de diferentes versões, com o intuito de averiguar a compatibilidade, o desempenho e a sincronização do dispositivo com o servidor em versões distintas. Sendo assim, o teste foi realizado em aparelhos com versões entre 4.0 até a versão 7.0.

Esta abordagem ajudou a identificar diferentes tipos de problemas incomuns. Por exemplo, nas versões 4.0, o aplicativo executava sem problema, mas a partir das versões 5.0, problemas desde a instalação começaram a surgir, levando o sistema a passar por atualizações.

Autores como [Wang et al. 2014], mencionados nos Trabalhos Relacionados, Capítulo 2, relatam que disponibilizaram aparelhos que possuíam uma mesma configuração para os alunos utilizarem durante sua pesquisa. Isto facilitou na uniformidade dos dados coletados e no uso da aplicação.

Nesta pesquisa, os celulares utilizados para instalação do aplicativo foram dos próprios estudantes que participaram. Sendo assim, um dos maiores desafios, deste trabalho, foi projetar o aplicativo para funcionar uniformemente em diversos dispositivos com arquiteturas e versões distintas. Por não se comportarem da mesma maneira, a aplicação levou ao desenvolvimento de outras atualizações e, por sua vez, novos testes tiveram que ser feitos a fim de garantir a resolução dos problemas.

Com isso, durante a utilização do *Cookie* pelos estudantes, surgiram problemas, principalmente em aparelhos cujas versões eram superiores a 5.0, tais como:

- erro ao instalar o aplicativo;
- inconsistência na coleta de algum dos tipos de dados;
- problema ao enviar os dados coletados;

Alguns dos problemas puderam ser identificados, virtualmente, a partir do uso do *Cookie*, pelos estudantes, com a utilização do *Fabric*³, uma plataforma gratuita de desenvolvimento móvel do Google que disponibiliza relatórios de estabilidade, acompanhamento de uso e de instalação do aplicativo, em tempo real, além de permitir o gerenciamento de novas versões.

Apesar das atualizações feitas para ajustar os problemas mencionados anteriormente, surgiram limitações (i.e versão do sistema operacional e modelo da arquitetura) de alguns *smartphones* que não puderam ser solucionadas, pois dependiam do usuário atualizar o sistema ou trocar o aparelho, o que levou a algumas desistências dos estudantes em utilizar o aplicativo.

Uma vez que as estratégias para serem utilizadas na pesquisa ficaram prontas, o projeto foi submetido ao Comitê de Ética.

3.3 Comitê de Ética e Pesquisa

O Comitê de Ética e Pesquisa (CEP) tem como missão, proteger os seres humanos envolvidos, de tal forma a garantir que seus interesses serão protegidos acima dos interesses da ciência ou da sociedade [Freitas e Hossne 2009].

Criado em 2001, o Comitê de Ética e Pesquisa da UEFS considera que a pesquisa com seres humanos é aquela que “individual ou coletiva, tenha como participante o ser humano, em sua totalidade ou partes dele, e o envolva de forma direta ou indireta, incluindo o manejo de seus dados, informações ou materiais biológicos” (Res. 466/12, II.14)⁴.

Assim, percebendo a necessidade, para a concepção do trabalho, de se fazer a coleta dos dados dos estudantes como, localização, horário de chegada e saída da sala,

³https://get.fabric.io/?utm_campaign=fabric-marketing&utm_medium=natural

⁴<http://cepuefs.wixsite.com/cepuefs/sobre>. Acessado em 03 de dez de 2017

participação, nome do aplicativo usado, local que senta e, ao mesmo tempo, protegê-los, foi submetido o projeto cumprindo as normas estabelecidas pelo Comitê de Ética no sentido de assegurar o respeito pela identidade e dignidade humana, bem como garantir o anonimato dos estudantes, respeitando sua integridade intelectual, social e cultural.

Para participar da pesquisa, os estudantes precisaram estar de acordo com alguns critérios que foram exigidos para inclusão e exclusão, são eles:

- **Critérios de inclusão:** o aluno precisou ter maioria (a partir de 18 anos); Aceitar participar e assinar o Termo de Consentimento Livre e Esclarecido - TCLE; Possuir celular com plataforma *Android* (caso aceite instalar o aplicativo).
- **Critérios de exclusão:** alunos que não estivessem matriculados nas turmas (EXA 801-T02 Algoritmos e Programação I; EXA 805-T01 Algoritmos e Programação II, EXA 806-T01 e T02 Estrutura de Dados), além dos estudantes que não concordassem em participar da pesquisa ou menores de 18 anos de idade.

Além disso, o projeto provê benefícios, assim como possíveis riscos relacionados à pesquisa, os quais foram relatados para o Comitê de Ética e também nos documentos assinados pelos estudantes e professores envolvidos, conforme Apêndice A e Apêndice B.

Quanto aos benefícios inerentes à pesquisa, estão aqueles relacionados ao processo de ensino aprendizagem de maneira geral, tais como a participação dos estudantes na aula, a compreensão de que os fatores podem ter relação com seu desempenho, inferir sobre dados não aparentes em grandes volumes de dados, através da mineração de dados, diante dos aspectos observados e dos dados coletados, refletir sobre possíveis soluções que contribuam para o processo de ensino aprendizagem e para o desempenho dos estudantes.

Já em relação aos riscos, existe o possível desconforto do participante em ser observado e em utilizar o aplicativo. Neste caso, a qualquer momento o estudante poderia desistir da pesquisa sem prejuízo para ele e/ou para a disciplina observada. Além disso, caso o aluno não desejasse ou desistisse de participar do projeto, o mesmo não seria penalizado nas disciplinas.

O projeto de pesquisa desse trabalho foi submetido e aprovado pelo Comitê de Ética, conforme Parecer Consubstanciado⁵ número 2.594.639.

3.4 Cenário de Pesquisa

A pesquisa foi executada na Universidade Estadual de Feira de Santana - UEFS. Implantada no Portal do Sertão, a UEFS surgiu de estratégia governamental com

⁵https://www.dropbox.com/s/lmr9um0dti7ulxz/Parecer_Consubstanciado_CEP.pdf?dl=0

objetivo de levar a educação superior para o interior, que até então estava restrita à capital, Salvador⁶. Atualmente, oferta 28 cursos de graduação os quais encontram-se distribuídos em quatro áreas de conhecimento: Tecnologia e Ciências Exatas; Ciências Humanas e Filosofia; Letras e Artes e Ciências Naturais e da Saúde.

O curso de Engenharia da Computação, oferecido desde 2003, está lotado nos departamentos de Tecnologia e Ciências Exatas. Atualmente, oferta 40 vagas por semestre e possui uma carga horária total de 3.955h, as quais estão distribuídas entre componentes e atividades.

Na observação em sala de aula, 76 estudantes aceitaram participar, o que corresponde a (100%) da amostra. Esta amostra ficou distribuída nas disciplinas da seguinte maneira: Algoritmos e Programação I, 18 alunos (23,7%); Algoritmos e Programação II, 19 alunos (25%); Estrutura de Dados Turma 01, 20 alunos (26,3%) e; Estrutura de Dados Turma 02, 19 alunos (25%).

Já com o *Cookie*, 40 estudantes (53,6%) concordaram instalar o aplicativo no *smartphone*. Dessa forma, a amostra das disciplinas ficou da seguinte maneira: Algoritmos e Programação I, 07 alunos (17,5%); Algoritmos e Programação II, 12 alunos (30%); Estrutura de Dados Turma 01, 10 alunos (25%) e; Estrutura de Dados Turma 02, 11 alunos (27,5%).

Nesta pesquisa, o quantitativo das amostras estudadas são diferentes. Sendo assim, foi considerada a amostra de 76 estudantes para análise da variável participação (fator interno) e a amostra de 40 estudantes para as Mídias Sociais (fator externo).

Como o desempenho acadêmico se refere a nota/conceito, utilizou-se a definição da Resolução 46/2006 adotada pela Universidade. De acordo com o Artigo 54, parágrafo 1º da resolução, a Avaliação do Estudante é feita ao longo do período apresentando três Medidas Parciais Consolidadas (MPC), sendo medida de forma numérica de zero a dez [0,0 - 10,0]. Em relação a média, o Artigo 56 informa que o valor total da Média das 3 (três) Medidas Parciais Consolidadas (MMPC) será determinado pela média aritmética simples. O Artigo 57, por sua vez, versa sobre a Avaliação Final, dizendo no 1º parágrafo que será aprovado e dispensado de realizar avaliação final o estudante que obtiver MMPC igual ou superior a 7,0 (sete); e, no 2º parágrafo informa que será considerado aprovado o estudante que obtiver Média Final (MF) igual ou superior a 5,0. Dessa forma, a média final (MF) de quem não fez a avaliação final é igual à sua MMPC.

Os estudantes foram observados em três disciplinas, ao longo de um semestre. Como o desempenho acadêmico de cada um deles foi mensurado a partir de sua média Global (MG), considerou-se para o cálculo da Média Geral da Turma (MGT), a média final (MF) dos estudantes. A Média Geral da Turma foi utilizada como parâmetro para avaliar a relação entre o desempenho acadêmico dos estudantes e o número de participações em cada turma e também para avaliar a relação entre o desempenho e o uso das Redes Sociais.

⁶<http://www.uefs.br/modules/conteudo/conteudo.php?conteudo=12>

3.5 Apresentação para as Turmas

A apresentação do projeto para as turmas aconteceu no primeiro dia de aula de cada turma. Para isso, foi feita uma apresentação/reunião com os estudantes para explicar a pesquisa e o uso do aplicativo. A apresentação indicou claramente quais dados seriam coletados durante a pesquisa, conforme explicados na Seção 3.1 e Seção 3.2 e, como estes dados seriam utilizados para análise e apresentação dos resultados.

Além disso, todos os requisitos estabelecidos pelo Comitê de Ética também foram apresentados (i.e critérios de inclusão e exclusão, riscos e benefícios, etc.) para que o estudante tivesse conhecimento dos termos éticos aplicados sobre os dados coletados e sua participação na pesquisa. Após a apresentação e os esclarecimentos, o estudante indicou seu desejo em participar ou não da pesquisa, assinando o TCLE, documento fundamental da pesquisa com ética.

3.6 Análise e Processamento dos Dados

O presente trabalho pretende analisar o fator interno à sala de aula: i) a relação entre a participação do estudante na sala de aula e; ii) o fator externo, a relação entre o uso de redes sociais e o desempenho acadêmico do estudante.

Para isso, a técnica utilizada foi a de mineração de correlação de *Spearman*. Este coeficiente é encontrado em estudos voltados para a Mineração de Dados Educacionais, como por exemplo, o estudo feito por Júnior e Outros [Júnior et al. 2017] que encontrou correlação entre fatores externos à sala de aula (i.e. uso de laboratórios de computação, salas de leitura e biblioteca) e o desempenho acadêmico positivo dos estudantes.

Esta técnica, também conhecida como aprendizado Não-Supervisionado, tem como finalidade encontrar correlações lineares, positivas ou negativas, entre diferentes atributos [Fayyad et al. 1996] e [Baker et al. 2011]. Segundo Dancey e Reidy [Dancey e Reidy 2018], este índice é recomendado quando a amostra conta com um número pequeno de participantes.

O coeficiente de correlação de *Spearman*, segundo Siegel e Castellan Jr [Siegel e Castellan Jr 1975], foi introduzido por Spearman (1904) e é considerada a mais antiga estatística baseada em postos. Este coeficiente, de acordo com Dancey e Reidy [Dancey e Reidy 2018], permite refletir a intensidade e a direção entre um conjunto de variáveis. Os valores do coeficiente podem variar de -1 (negativo) a +1 (positivo) e 0 (zero), no qual -1 apresenta existência de correlação forte negativa (e.g. alto uso das mídias sociais e baixo desempenho), +1 indica que há correlação forte positiva (e.g. alto uso das mídias sociais e o alto desempenho), e quando for 0 (zero) não apresenta existência de correlação.

Para este trabalho, foi considerado como significativo o resultado em que o seu valor, conhecido como Alfa (α), probabilidade utilizada como ponto de corte, foi igual ou

inferior a 0,05, Dancey e Reidy [Dancey e Reidy 2018]. Quanto a classificação, os autores explicam que quando a correlação apresentar valores entre 0,1 e 0,3 são considerados fracos; entre 0,4 e 0,6 pode ser considerado moderado ou médio, e entre 0,7 e 1,0 são entendidos como fortes. Porém é importante salientar que quanto mais próximo de 1, mais forte é a correlação [Dancey e Reidy 2018].

As variáveis consideradas como fatores interno e externo para a análise foram: i) interno: a taxa de participação nas aulas por aula, pelo estudante; ii) externo: nome do aplicativo em uso, data da coleta e; iii) o desempenho acadêmico, a Média Global (GB) gerada pelas três notas do semestre a partir das avaliações realizadas pelas disciplinas. O processo da coleta dos dados bem como a aplicação do KDD encontram-se explicadas no Capítulo 4.

Capítulo 4

Coleta e Análise de Dados

Este capítulo descreve a etapa da coleta dos dados, que foi feita por meio da observação e através do uso do aplicativo *Cookie*, pelos estudantes. Além disso, são apresentadas as etapas do KDD utilizadas na mineração de dados educacionais.

4.1 Coleta dos Dados

A coleta de dados, feita por meio da técnica de observação, foi realizada no período de 1 semestre, ou seja, 4 meses letivos. Durante essa fase, a pesquisadora acompanhou todas as aulas das disciplinas (Algoritmos e Programação I, Algoritmos e Programação II, Estrutura de Dados T01 e Estrutura de Dados T02), anotando em uma planilha todas as vezes que um estudante participava/interagia com o professor ou com outro colega, durante a aula.

Já a coleta de dados, feita através da utilização do *Cookie*, foi realizada no período de 2 meses, durante as aulas. Nesta etapa, os estudantes instalaram o aplicativo em seus *smartphones*, a partir de um link disponibilizado. Após a instalação, a detecção dos dados aconteceu de forma automática, sem a interferência do estudante.

Desta forma, não foi necessário que o estudante tivesse um pacote de dados para utilizar o aplicativo. Quando o celular era conectado a uma rede *Wi-Fi*, os dados eram enviados, automaticamente, para uma base de dados em um servidor, localizado no LABOTEC III - UEFS. Dos *smartphones* dos estudantes, foram coletados 80MB de dados.

4.2 Aplicação da Mineração de Dados

Nesta etapa, foram utilizados os dados extraídos da planilha (Observação) e do banco de dados (*Cookie*). Este processo foi adaptado do KDD, Figura 2.2, com o intuito de alcançar os objetos da pesquisa.

Para que este processo fosse alcançado, foram feitos os seguintes passos: (a) análise e seleção dos dados; (b) pré-processamento e transformação dos dados; (c) aplicação da técnica de mineração de correlação e; (e) interpretação dos dados, todos eles explicados nas seções seguintes.

Todas as etapas do KDD foram desenvolvidas em Python, utilizando as bibliotecas: (a) *Pandas*:¹ fornece estruturas de alto desempenho para análise de dados; (b) *NumPy*:² um pacote utilizado em computação científica e; (c) *Scipy.stats*:³ uma biblioteca que possui funções estatísticas.

• Análise e Seleção de Dados

Nessa fase, foi feita a análise e seleção de todos os atributos coletados nas duas bases obtidas com a coleta (Observação e *Cookie*). Esta foi uma etapa importante, pois ajudou a entender a estrutura dos dados, para que fosse possível extrair as informações necessárias para o estudo. A Tabela 4.1 apresenta os atributos selecionados com suas respectivas descrições. Estes atributos são referentes a cada estudante, por disciplina.

Tabela 4.1: Junção das variáveis coletadas

Variáveis	Coleta	Descrição
Participação 1	Observação	Número de participações na 1ª unidade
Participação 2	Observação	Número de participações na 2ª unidade
Participação 3	Observação	Número de participações na 3ª unidade
Nota 1	Observação	Nota da 1ª unidade
Nota 2	Observação	Nota da 2ª unidade
Nota 3	Observação	Nota da 3ª unidade
Media	Observação	Média das notas no semestre
Media Final	Observação	Média Final do semestre
Imei	Obs/Cookie	Identificação do celular do aluno
Aplicativo Aberto	Cookie	Nome do aplicativo
Data_Coleta	Cookie	Data da coleta

Fonte: Próprio Autor.

Na variável Participação 1, foram registradas todas as vezes que o estudante participou da aula, na primeira unidade. Na segunda unidade registraram-se as intervenções dos estudantes na Participação 2 e, na Participação 3, todas as participações referentes à terceira unidade. Foi considerada participação, neste trabalho, qualquer intervenção feita por um estudante em relação ao professor, seja ela uma pergunta ou um comentário. Foi considerada também toda pergunta feita por um estudante e respondida por outro colega, na tentativa de esclarecer ou ajudar com alguma dúvida.

¹<https://pandas.pydata.org/>

²<https://www.numpy.org/>

³<https://docs.scipy.org/doc/scipy/reference/stats.html#module-scipy.stats>

A Nota 1 é a variável na qual está armazenada a nota do estudante na primeira unidade do semestre. A Nota 2 armazena a nota da segunda unidade, e a Nota 3 armazena a nota da terceira unidade.

A variável Media armazena o total obtido pelas (03) três notas do semestre, que é calculado pela média aritmética simples.

A Media Final é calculada após a Média, ou seja, é obtida por aquele estudante que não alcançou a nota 7,0 (aprovação) e, para aqueles que obtiveram aprovação, a média final é igual a média.

O atributo Imei guarda o número de registro do *Smartphone* do estudante. Esse identificador é único e também foi cadastrado na planilha utilizada na observação, o que ajudou no momento de fazer a junção com a tabela do *Cookie*.

A variável AplicativoAberto registrou o nome do aplicativo que estava em uso no momento da coleta. E a Data_Coleta coletou o dia, mês, ano, além da hora, minuto e segundo da coleta. É importante salientar que o aplicativo faz a coleta por minuto, sem a interferência do estudante.

• Pré-Processamento e transformação

Nesta etapa, foi feita a limpeza dos dados inconsistentes e a padronização dos dados em formato adequado para aplicação da mineração. Como existem dois tipos de coleta, os dados ficaram armazenados em duas tabelas distintas, uma para a análise dos dados coletados com a observação e outra para os dados coletados com o *Cookie*.

Dentre as variáveis coletadas durante a observação, a participação foi selecionada para a análise. Esta seleção foi feita com base nos estudos de Kuh [Kuh 2005, Kuh 2009] e Fior e Mercuri [Fior e Mercuri 2018] que afirmam que a participação do estudante em tarefas distintas, seja dentro ou fora da sala, bem como o tempo e a energia empenhados nas atividades acadêmicas é um fator primordial para o sucesso acadêmico. As variáveis utilizadas na etapa da mineração encontram-se na Tabela 4.2.

Tabela 4.2: Variáveis selecionadas na Observação

Variáveis	Tipo	Descrição
TaxaParticipação	Numérico	Total de participações no semestre
MediaFinal	Numérico	Média Final do semestre

Fonte: Próprio Autor.

As participações dos estudantes, coletadas durante o semestre, conforme apresentada na Tabela 4.1, gerou o atributo TaxaParticipação. Para encontrar esta taxa, a soma das três participações foi dividida pela média de estudantes da turma (número de estudantes no início e no final do semestre). Por exemplo, em uma turma que começou com 25 estudantes e finalizou com 20, a média aritmética simples da turma

é 22,5. Nessa mesma turma, o total de participações foi de $800 / 22,5 = 35,5$, ou seja uma taxa de 35,5 participações por turma, ao longo do semestre.

Já com o *Cookie*, dentre as variáveis coletadas, a selecionada para análise foi o AplicativoAberto. Apesar de existirem vários tipos de aplicativos utilizados pelos estudantes, as Mídias Sociais (*WhatsApp*, *Facebook*, *Instagram* e *YouTube*) foram as escolhidas para análise, pois segundo dados do Global Digital Report ⁴, estas foram as plataformas de mídias sociais mais ativas em 2017, utilizadas pelos brasileiros. A Tabela 4.3 apresenta as variáveis utilizadas na mineração.

Tabela 4.3: Variáveis selecionadas do *Cookie*

Variáveis	Tipo	Descrição
AplicativoAberto	Numérico	Nome do aplicativo
MediaFinal	Numérico	Média Final do semestre

Fonte: Próprio Autor.

Na variável AplicativoAberto, foi coletada a quantidade de vezes que cada aplicativo foi acessado durante o semestre. Assim, foi encontrada a média de acesso por estudante de cada turma, a qual foi utilizada na análise.

• Mineração de Dados

Para encontrar os resultados com os dados coletados por meio da observação, foi utilizado o índice de correlação de *Spearman* com as variáveis TaxaParticipação e a MédiaFinal, apresentadas na Tabela 4.2.

Para os resultados com os dados do *Cookie*, o índice de correlação de *Spearman* foi utilizado com as variáveis AplicativoAberto, sendo escolhidas as mídias sociais (*WhatsApp*, *Facebook*, *Instagram* e *YouTube*) e a MédiaFinal, ver Tabela 4.3 .

O objetivo é identificar se existe relação entre os fatores interno (participação) e externo (Mídias Sociais) o desempenho acadêmico. Esta operação foi realizada com cada turma, que foi acompanhada no estudo. Os resultados gerados, bem como a análise e interpretação dos dados encontram-se no Capítulo 5.

⁴<https://www.slideshare.net/wearesocial/digital-in-2018-in-southern-america-part-1-north-86863727>>. Acesso em abril, 2019.

Capítulo 5

Resultados

Este capítulo apresenta na Seção 5.1 os resultados separados por turma (Algoritmos e Programação I, Algoritmos e Programação II e Estrutura de Dados turmas T01 e T02) e, em seguida, é apresentado o resultado da junção dos dados coletados de todas as turmas. O mesmo processo é feito com a Seção 5.2. O intuito é verificar a diferença entre os resultados e se estes podem confirmar ou refutar as hipóteses de pesquisa.

5.1 Estudo experimental dos fatores internos

Os resultados encontrados sobre a participação do estudante, se encontram organizados em dois tópicos. No primeiro momento, é apresentada uma análise descritiva dos dados sobre o desempenho acadêmico dos estudantes ao final do semestre, e sobre a participação dos estudantes nas aulas (Seção 5.1 e Tabela 5.1). Segundo afirmação dos autores Reis e Reis [Reis e Reis 2002], a análise descritiva é utilizada a fim de organizar, resumir e descrever os aspectos importantes para comparar as características fundamentais entre os conjuntos de dados observados. Para compor os intervalos de desempenho no semestre, foi utilizado o percentil, culminado na divisão de 3 (três) níveis de desempenho para cada disciplina observada.

No segundo momento, são apresentados os resultados da relação entre as duas variáveis (participação e o desempenho acadêmico), verificada por meio do índice de correlação de *Spearman* (Tabela 5.1).

- **Desempenho Acadêmico**

A Tabela 5.1 apresenta o resultado da distribuição dos estudantes, nas disciplinas estudadas, em conformidade com o desempenho acadêmico.

No nível de baixo desempenho, encontra-se a maioria dos estudantes, somando um total de 50%, dos quais 26,3% aparecem na disciplina de Estrutura de Dados T01 com notas mínima e máxima de 0,4 e 9,5, respectivamente e Média Global 4,6 e,

23,7% estão na disciplina de Algoritmos e Programação I, com nota mínima 1,0 e a máxima 9,2 obtendo média de 4,1.

Tabela 5.1: Distribuição dos estudantes nas disciplinas em relação ao nível de desempenho acadêmico

	Disciplina	n	Média	Desvio Padrão	Min	Max
Baixo Desempenho	Estrutura de Dados T01	20	4,6	2,833	0,4	9,5
	Alg. e Programação I	18	4,1	2,759	1,0	9,2
Desempenho Mediano	Alg. e Programação II	19	6,4	3,215	0,7	10,0
Melhor Desempenho	Estrutura de Dados T02	19	7,1	2,698	0,5	10,0

Na disciplina Algoritmos e Programação II, 25% dos estudantes se encontram com desempenho mediano. Com notas mínima e máxima 0,7 e 10,0, e média 6,4. No nível do melhor desempenho está a disciplina Estrutura de Dados T02, 25% dos estudantes. Nesse nível, os estudantes apresentam a nota mínima 0,5 e a máxima 10,0 e média 7,1.

Os dados apresentados na disciplina de Algoritmos e Programação I pode ser explicado pelo fato da turma ser de calouros e os estudantes vivenciarem seu primeiro contato com uma linguagem de programação.

A Tabela Tabela 5.2, apresenta o quantitativo de estudantes e o percentual de estudantes que se encontram em cada faixa de notas.

Tabela 5.2: Distribuição do desempenho dos estudantes no semestre

	Alg. e Prog. I	Alg. e Prog. II	Estrutura de Dados T01	Estrutura de Dados T02	Frequência Total	%
[9,0 - 10,0]	1	3	1	5	10	13,2
[7,0 - 8,99]	4	10	5	9	28	36,8
[5,0 - 6,99]	3	0	4	2	9	11,8
[3,0 - 4,99]	1	1	2	1	5	6,6
[0,0 - 2,99]	9	5	8	2	24	31,6

Os dados dos estudantes se apresentam em dois grupos: estudantes na faixa de aprovação [5,0 - 10,0] e aqueles na faixa de reprovação [0,0 - 4,99]. Observa-se, pela Tabela 5.2, que na amostra estudada, há uma boa parcela de estudantes que representa o grupo de reprovação 38,2%. Na faixa de aprovação, 11,8% foram aprovados após a avaliação final, e 50%, sem avaliação final, com média superior ou igual a 7,0. Assim, é notável uma frequência maior de estudantes com desempenho variando entre 7,0 e 8,9, uma das faixas de aprovação adotada pela Universidade.

• Participação dos Estudantes nas Aulas

A Tabela 5.3 apresenta a quantidade de estudantes no início (Turma inicial) e no final do semestre (Turma Final), de cada uma das turmas; a quantidade média de estudantes (Média), e, por fim, a taxa de participação por aluno durante o semestre. Observa-se, pelos dados, que a maior taxa de participações localiza-se na disciplina de Algoritmos e Programação I, com 51,7. Em seguida, aparece Algoritmos e Programação II com 30,7. A disciplina de Estrutura de Dados T02 apresenta uma taxa de 21,6, e por último, Estrutura de Dados T01 com 12,4.

Tabela 5.3: Distribuição da participação dos estudantes

	Turma Inicial	Turma Final	Média	Taxa de participação
Alg. e Programação I	24	18	21,0	51,7
Alg. e Programação II	24	19	21,5	30,7
Estrutura de Dados T01	24	20	22,0	12,4
Estrutura de Dados T02	24	19	21,5	21,6

Há uma taxa alta de participação dos estudantes nas aulas de Algoritmos e Programação I, participação que pode ser explicado pelo fato de que esta é uma turma formada apenas por calouros, enquanto as demais por veteranos e/ou concluintes. Este resultado vai de encontro com a afirmação de [Fior e Mercuri 2018], de que a participação acadêmica pode variar em função do grau do estudante (i.e. calouros, concluintes), “a participação acadêmica dos estudantes é distinta entre calouros e concluinte” [Fior e Mercuri 2018, p.88].

• A Participação nas aulas e o Desempenho Acadêmico

A fim de investigar a relação entre a participação nas aulas e o desempenho acadêmico entre os estudantes do curso de Engenharia da Computação, aplicou-se o índice de correlação de *Spearman*.

Ao analisar o coeficiente de correlação de *Spearman* entre a participação do estudante e o desempenho acadêmico em **Algoritmos e Programação I** foi encontrada uma

relação de $r=-0,60$; $p=0,399$. Isto significa dizer que para a amostra estudada, os dados evidenciam uma relação negativa significativa, apesar de ter apresentado um valor moderado. O resultado entre a participação e o desempenho, encontrado na disciplina de **Algoritmos e Programação II**, $r=0,32$; $p=0,434$ demonstra uma correlação estatisticamente significativa, embora a amostra tenha apresentado um valor positivo fraco.

Para a disciplina de **Estrutura de Dados T01**, os resultados apresentaram uma relação fraca $r=0,30$; $p=0,623$, o que significa dizer que para esta amostra as variáveis caminharam independentes, não apresentando resultados estatisticamente significativos. Em **Estrutura de Dados T02**, a relação encontrada foi de $r=0,07$; $p=0,847$ entre as variáveis estudadas, o que também significa dizer que existe uma relação fraca, não significativa.

Ao analisar o fator interno (participação), na disciplina de Algoritmos e Programação I, os resultados se opõem aos estudos de Costa e Outros [Costa et al. 2014], pois, para a relação entre a participação e o desempenho do estudante, foi encontrada uma correlação moderada negativa, o que significa dizer que estudantes com maior número de participação apresentaram baixo desempenho. Para essa amostra, os dados revelam que apesar de apresentar valor significativo estatisticamente, a relação não é clara e constante para ser conclusiva. Na disciplina de Algoritmos e Programação II, os resultados apresentaram uma correlação fraca, ou seja, não existe relação entre a participação e o desempenho. Em Estrutura de Dados, turmas T01 e T02, os resultados também apresentaram uma correlação fraca, indicando que diferente de Fior e Outros [Fior et al. 2013], para estas disciplinas, não foi encontrada uma relação entre a participação e o desempenho, refutando a hipótese.

Um fator que pode ter influenciado neste resultado, é que ao separar as turmas por disciplina, o número de participantes cai muito, diminuindo também a quantidade de dados coletados. Assim, optou-se em fazer a análise da relação de todos os dados juntos, ou seja, unindo os dados de todas as turmas. Neste caso, foi encontrada uma relação de $r=-0,562$; $p<0,001$. O resultado indica que apesar da amostra apresentar valor estatisticamente significativo a sua relação é moderada, o que significa dizer que a amostra não apresenta possibilidades de descrever uma relação. Este resultado é semelhante ao da disciplina de Algoritmos e Programação I.

Embora alguns estudos relatem a existência de relação entre a participação e o desempenho acadêmico, outros fatores podem ter influenciado nos resultados deste estudo. Por exemplo, o uso das notas para medir o desempenho, uma vez que outras variáveis têm influências sobre elas, como motivação dos estudantes, conteúdos abordados e interesses dos estudantes, horas de dedicação aos estudos entre outros.

5.2 Estudo experimental dos fatores externos

Os resultados encontrados com o uso do *Cookie* também estão organizados em dois tópicos. No primeiro momento, é apresentada uma análise descritiva dos dados

sobre o desempenho acadêmico dos estudantes ao final do semestre e sobre o uso das Mídias Sociais no *Smartphone* (Seção 5.2 e Tabela 5.2). No segundo momento, são apresentados os resultados da relação entre as duas variáveis por meio do índice de correlação de *Spearman* (Tabela 5.2).

• Desempenho Acadêmico

A Tabela 5.4 apresenta o resultado dos estudantes das quatro turmas em conformidade com o desempenho da Média Global, durante o semestre. Como pode ser observado, 25% dos estudantes se encontram na disciplina Estrutura de Dados - T01, apresentando um baixo desempenho. Isso significa que as notas mínima e máxima dos estudantes, nesta disciplina, foram de 0,4 e 7,4, respectivamente, obtendo uma Média Global 4,9.

Tabela 5.4: Distribuição das turmas em relação ao desempenho acadêmico nas disciplinas.

	Disciplina	n	Média	Desvio Padrão	Min	Max
Baixo Desempenho	Estrutura de Dados T01	10	4,9	3,133	0,4	7,4
Desempenho Mediano	Alg. e Programação I	7	7,3	1,104	3,9	9,2
	Alg. e Programação II	12	7,8	1,605	1,3	9,6
Alto Desempenho	Estrutura de Dados T02	11	8,7	1,629	3,0	10,0

No nível de desempenho mediano, encontra-se a maioria dos estudantes, um total de 47,5%, sendo eles: i) 17,5% dos estudantes da disciplina Algoritmos e Programação I com notas mínima de 3,9 e máxima de 9,2, com média 7,3, e; ii) 30% dos estudantes na disciplina de Algoritmos e Programação II, com nota mínima 1,3 e a máxima 9,6 e média 7,8. O melhor desempenho está localizado na disciplina de Estrutura de Dados - T02, com notas mínima e máxima de 3,0 e 10,0, respectivamente, obtendo uma média de 8,7.

Na Tabela 5.5, é apresentado o desempenho por disciplina durante o semestre, representado pela Média Global dos estudantes, de acordo com os intervalos das notas, considerados pela Resolução 46/2006, citada anteriormente.

Conforme a Tabela 5.5, nota-se que os dados dos estudantes se encontram discriminados em dois grupos: aprovados e reprovados. Com isso, é possível concluir que a amostra estudada apresenta características homogêneas, pois os estudantes que representam o grupo de reprovação correspondem a 22,5%, enquanto que o grupo dos estudantes que estão na faixa de aprovação apresenta os percentuais de 12,5%

Tabela 5.5: Distribuição do desempenho dos estudantes no semestre.

	Alg. e Prog. I	Alg. e Prog. II	Estrutura de Dados T01	Estrutura de Dados T02	Frequência Total	%
[9,0 - 10,0]	1	2	1	2	6	15
[7,0 - 8,99]	3	7	4	6	20	50
[5,0 - 6,99]	2	0	1	2	5	12,5
[3,0 - 4,99]	1	1	1	1	4	10
[0,0 - 2,99]	0	2	3	0	5	12,5

e 65%. Assim, é notável a alta frequência de estudantes com desempenho variando entre 7,0 e 8,99, uma das faixas de aprovação.

• Uso das Mídias Sociais

Dentre os aplicativos (i.e., *Chrome*, fotos, câmera, *e-mail*, jogos) coletados através do *Cookie*, foram selecionados para análise apenas aqueles considerados Mídias Sociais, (i.e. *WhatsApp*, *YouTube*, *Instagram* e *Facebook*). Na Tabela 5.6, apresenta-se a quantidade de utilização de cada aplicativo, neste período, de acordo com cada turma.

Tabela 5.6: Distribuição da utilização dos aplicativos.

	Alg. e Prog. I	Alg. e Prog. II	Est. de Dados T01	Est. de Dados T02	Frequência Total	%
WhatsApp	81	6999	909	6487	14476	75,02
YouTube	6	1012	41	238	1297	6,72
Instagram	15	408	21	454	906	4,70
Facebook	0	281	95	402	778	4,03

Observando os dados da Tabela 5.6, de maneira geral, nota-se uma grande diferença entre a utilização dos aplicativos de Mídias Sociais. O *WhatsApp* aparece liderando com 75,02%, enquanto o *Instagram* e o *Facebook*, com apenas 4%, apresentando os percentuais mais baixos. A utilização do *YouTube* aparece com 6,72%.

Dessa forma, é possível perceber que o *WhatsApp* foi a mídia mais utilizada pelos estudantes. Este alto número de interação é justificado pelo fato do seu mecanismo de troca de mensagens e arquivos acontecer em tempo real. De acordo com Smith,

[Smith 2019] o *WhatsApp* ultrapassou 1,5 bilhões de usuários em 2018, chegando a 65 bilhões de envio de mensagens, mais de 4,5 bilhões de fotos e 1 bilhão de vídeos enviados diariamente.

Outro fator a ser destacado é que quando os usuários acessam o *Facebook*, via *browser Chrome*, por exemplo o *Cookie* armazena o nome *Chrome* e não *Facebook*. Isto pode explicar a grande diferença de uso entre os aplicativos, visto que muitos usuários utilizam o YouTube e o *Facebook* via *browser*.

• O uso de Mídias Sociais e o Desempenho Acadêmico

A análise foi feita por disciplina, com os aplicativos de Mídias Sociais: *WhatsApp*, YouTube, *Instagram* e *Facebook*. Para isto, utilizou-se a média de acesso que cada aplicativo foi acessado pelo estudante, conforme apresentado na Tabela 5.6.

Ao analisar a correlação entre o uso dos aplicativos mencionados anteriormente, na disciplina de **Algoritmos e Programação I**, não foi possível encontrar valores significativos do índice de *Spearman*. Ainda assim, é possível destacar o *WhatsApp* com maior frequência de uso (Tabela 5.6).

Com a disciplina **Algoritmos e Programação II**, a relação entre o uso do *WhatsApp* o desempenho foi de ($r = -0,030$; $p = 0,953$), evidenciando uma relação fraca não significativa. Com o YouTube, os dados também revelaram uma relação fraca ($r = 0,234$; $p = 0,654$), o que significa dizer, ao menos para essa amostra, que as variáveis caminharam independentes e não apresentaram resultados significativos. No que diz respeito ao uso *Facebook*, foi encontrada uma relação forte e valor estatisticamente significativa ($r = 0,866$; $p = 0,333$). Isso significa dizer que, os estudantes que utilizaram muito o *Facebook* apresentaram bom desempenho acadêmico. Com o *Instagram*, os resultados indicaram uma relação forte negativa significativa ($r = -0,866$; $p = 0,333$), indicando que quanto maior a utilização da mídia, menor é o desempenho dos estudantes.

Na disciplina **Estrutura de Dados T01**, na relação entre o uso do *WhatsApp*, do *Facebook* e do *Instagram* com o desempenho ($r = 0,999$; $p < 0,001$), os resultados foram semelhantes, indicando uma relação forte positiva significativa, ou seja, os estudantes que apresentaram melhor desempenho acessaram mais vezes as Mídias Sociais. Já com o YouTube, a amostra não apresentou resultados com possibilidades para descrever uma relação.

Em **Estrutura de Dados T02**, os resultados revelaram uma relação fraca ($r = 0,231$; $p = 0,658$) entre o uso do *WhatsApp* e o desempenho. Já com o YouTube, houve uma correlação significativa, embora tenha sido apresentado um valor moderado ($r = 0,573$; $p = 0,234$). Para o uso do *Instagram*, foi encontrada uma relação fraca, não significativa ($r = 0,399$; $p = 0,600$). E com o uso do *Facebook* ($r = 0,316$; $p = 0,683$), os resultados também apresentaram uma relação fraca. Nesta turma, os resultados não foram conclusivos para o uso das Mídias Sociais em relação ao desempenho.

Na disciplina de Algoritmos e Programação I, não foram encontrados valores significativos para análise. Mesmo assim, é possível destacar uma maior frequência com

o uso do *WhatsApp*. O que pode ter influenciado neste resultado é o número de estudantes, apenas 07, Seção 3.4, concordaram em utilizar o *Cookie*.

Na disciplina de Algoritmos e Programação II, foi encontrada uma relação forte positiva com o *Facebook*, o que significa dizer que os estudantes que utilizaram muito a mídia, tiveram bom desempenho. Para o *Instagram*, foi encontrada uma relação forte negativa, indicando que, nesta disciplina, quanto maior foi o uso do *Instagram*, menor foi o desempenho do estudante. Este resultado é semelhante ao estudo de Kirschner e Karpinski [Kirschner e Karpinski 2010], porém a relação negativa significativa, neste estudo, foi encontrada entre o *Instagram* e o desempenho, e forte significativa, com o *Facebook*. Com relação ao uso do *WhatsApp* e do YouTube, os resultados apresentaram uma relação fraca, ou seja, não existe correlação entre estas mídias sociais e o desempenho.

Em Estrutura de Dados T01, foi encontrada uma relação positiva forte com o *WhatsApp*, o *Facebook* e o *Instagram*, o que significa dizer que os estudantes que utilizaram muito estas mídias tiveram bom desempenho.

Na disciplina Estrutura de Dados T02, os dados apresentaram uma correlação fraca com *WhatsApp*, e moderada com o YouTube, porém essa relação não é clara e constante para ser conclusiva. Já o uso do *Instagram* e do *Facebook* apresenta uma relação fraca, indicando que para esta disciplina os resultados não há relação entre o uso das Mídias Sociais e o desempenho. Nesta disciplina, os resultados corroboram com os estudos de Rangel e Miranda [Rangel e Miranda 2016], Silva e Outros [Silva et al. 2012], Campos e Outros [Campos et al. 2012], que afirmam que o uso das Mídias Sociais não tem relação com o rendimento acadêmico.

Nesse caso, para o fator externo mídias sociais, obtido a partir de cada turma, os dados evidenciaram uma relação forte significativa entre o uso do *Facebook* e o desempenho. E apresentou uma relação negativa significativa entre o uso do *Instagram* e o desempenho, confirmando a hipótese de que existe relação entre o uso de Mídias Sociais e o desempenho.

Assim como na Seção 5.1, também foi realizada uma análise com a junção dos dados de todas as turmas, os quais foram obtidos os seguintes resultados. A relação entre o uso do *WhatsApp* e o desempenho foi de ($r=0,627$; $p=0,051$), evidenciando uma correlação moderada significativa. Com o *YouTube*, os dados mostraram uma relação moderada não significativa ($r=0,611$, $p=0,080$). No que diz respeito ao *Facebook* a amostra apresentou uma relação moderada estatisticamente não significativa ($r=0,698$; $p=0,122$). E com o *Instagram* ($r=0,781$; $p=0,0377$), os dados revelaram uma relação forte positiva significativa, o que significa dizer que aqueles estudantes que utilizaram muito esta mídia social apresentaram um bom desempenho acadêmico.

Diferente dos resultados obtidos por turma, na junção dos dados fica evidente uma relação forte positiva e estatisticamente significativa entre o uso do *Instagram* e o desempenho.

Capítulo 6

Considerações Finais

Este trabalho buscou investigar dados que se relacionam com o desempenho acadêmico do estudante, a partir de dados coletados dentro e fora da aula. O objetivo principal foi aplicar técnica de mineração de dados para encontrar correlações entre os fatores interno e externo, mais especificamente a participação em sala de aula (interno) e o uso de redes sociais (externo) e o desempenho acadêmico de estudantes do curso de Engenharia da Computação na Universidade Estadual de Feira de Santana.

Nesse sentido, pode-se afirmar que o objetivo da pesquisa foi cumprido. Para isso, aplicou-se a mineração de dados educacionais seguindo as etapas do KDD, o que possibilitou, através da mineração de correlação, investigar a existência de relações nos dados coletados por meio da observação em sala de aula, e o uso do *Cookie*, pelos estudantes.

Para o desempenho do estudante, foi utilizada a Média Global das disciplinas estudadas, gerada a partir das três notas do semestre. Os estudantes apresentaram um rendimento satisfatório, pois a maioria se encontra no intervalo das notas entre 7,0 e 8,99, uma das faixas de aprovação adotada pela Universidade. Por outro lado, existe uma frequência considerável de alunos na faixa de reprovação, uma média de 36,6% do total. Esse valor, apesar de estar abaixo, quando comparado com os dados apresentados por [Bittencourt et al. 2013], nos quais a reprovação se encontrava entre 50% e 75% nas disciplinas iniciais de programação entre 2009 e 2011, há uma possibilidade que volte a se modificar (crescer ou decrescer) em decorrência de outros fatores que determinam o desempenho, conforme visto no decorrer deste trabalho e na literatura.

Como o desempenho acadêmico, nesse trabalho, é entendido como as notas do estudante, muitas são as variáveis que impactam o desempenho do estudante ao longo de sua formação acadêmica superior. Assim, buscou-se escolher algumas variáveis, para entendê-las isoladamente, para em trabalhos futuros, fazer correlações, inferências sobre elas, referências cruzadas.

Com relação ao uso das Mídias Sociais (fator externo), um fator que merece importância é sobre a alta frequência da utilização das mesmas (*WhatsApp*, YouTube, *Instagram* e *Facebook*). Os resultados corroboram com os dados apresentados pelo Global Digital Report (2019), que apresentam estas como as Mídias Sociais mais utilizadas pelos brasileiros atualmente, com alteração apenas da ordem de maior utilização. Naquela pesquisa, o YouTube aparecia como o mais utilizado e, neste resultado, o *WhatsApp* é o primeiro deles.

Este trabalho, fez também, a coleta de outras variáveis que não foram utilizadas na pesquisa, e que poderiam ajudar a entender os resultados encontrados como, o local que o aluno senta (frente, meio e fundo), a localização geográfica (latitude e longitude) e o estado do celular (parado/movimento).

Este estudo contribui para a área da educação, pois os resultados podem colaborar nas causas que incidem sobre a retenção e evasão, como problemas de aprendizagem, fatores externos que levam à procrastinação, práticas pedagógicas, dificuldades nas disciplinas, [MEC 2014], entre outros, possibilitando que o professor possa utilizar e refletir sobre as informações geradas a partir dos resultados em seus ambientes acadêmicos, como forma de adaptar estratégias em sala de aula, para o envolvimento dos estudantes. Além disso, o estudo apresenta o potencial de uso da mineração de dados para gerar resultados e que o ciclo do KDD tem sua importância nesse processo. O uso da correlação é importante pois comprova a existência ou não de relação entre as variáveis, sendo utilizado em outros trabalhos na área de mineração de dados educacionais.

6.1 Pesquisas Futuras

Os estudos mostraram que outros fatores que não sejam a frequência de uso das redes, como horas de estudo, a participação nas aulas, a dedicação com a disciplina, estudo em outros ambientes (i.e bibliotecas), o turno da aula, os estudos em grupos entre outros, podem ser considerados para uma análise futura das suas relações e impactos no desempenho discente.

Para trabalhos futuros, a sugestão é coletar outros dados (i.e assunto, atividades, prática) por meio da observação e/ou até mesmo aplicar questionário, a fim de verificar outras variáveis, como, por exemplo, a motivação do aluno na disciplina ou no curso, para identificar fatores outros que possam influenciar o baixo ou o alto desempenho acadêmico. Outra questão importante para ser trabalhada futuramente é não só enumerar a quantidade de participações/intervenções, como foi feito neste experimento, como também observar a qualidade destas. Nem sempre, se colocar em sala de aula representa algo significativo, pois pode ser, por exemplo, uma intervenção que não tenha correlação com o assunto/aula trabalhados. Sugere-se, também, ampliar esse trabalho para uma pesquisa quali-quantitativa, com associação dos números obtidos com as percepções e respostas dos participantes envolvidos.

Dessa forma, é importante perceber a relação entre a participação/intervenção discente com a qualidade destas falas/perguntas e com os conteúdos abordados em aula, além também do *feedback* do professor e o quanto isso impactou na evolução da aprendizagem. Geralmente, os trabalhos que avaliam o desempenho acadêmico, têm muito a perspectiva do discente, e pouco a percepção do professor em relação a cada estudante. Talvez isso seja em função das salas com grande quantidade de estudantes, mas seria algo interessante, para perceber a avaliação processual, e não apenas somativa (notas). Saber como o aluno está, como estava e o quanto evoluiu. A nota nem sempre reflete o aprendizado, mas é o parâmetro, na educação nacional usado para aprovação/reprovação.

Com relação a coleta de dados através do *smartphone* dos estudantes, sugere-se a melhoria do aplicativo desenvolvido, pois, para este trabalho, o *Cookie* só pode ser utilizado em plataformas Android, o que gerou ausência de alguns voluntários. Além disso, outras variáveis podem ser acrescentadas ao aplicativo como os assuntos das disciplinas, a metodologia trabalhada, o motivo do uso da rede social (i.e., fins acadêmicos, estudos, etc.), dentre outras, para que seja possível identificar outros fatores que impactam ou tenham relação no desempenho dos estudantes. Essas novas variáveis podem, não só ampliar a base pesquisada, como também, de maneira mais assertiva, trazer resultados que justifiquem o uso acadêmico ou não das redes sociais.

Referências

- [Albuquerque 2016] Albuquerque, T. (2016). Do abandono à permanência num curso de ensino superior. *Sísifo*, (7):19–28.
- [Algarni 2016] Algarni, A. (2016). Data mining in education. *IJACSA) International Journal of Advanced Computer Science and Applications*, 7(6).
- [Amaral 2016] Amaral, F. (2016). *Aprenda mineração de dados: teoria e prática*, volume 1. Alta Books Editora.
- [Baker et al. 2010] Baker, R. et al. (2010). Data mining for education. *International encyclopedia of education*, 7(3):112–118.
- [Baker et al. 2011] Baker, R., Isotani, S., e Carvalho, A. (2011). Mineração de dados educacionais: Oportunidades para o brasil. *Brazilian Journal of Computers in Education*, 19(02):03.
- [Bittencourt et al. 2013] Bittencourt, R. A., Rocha, A. S., Santana, B. L., Santana, C. S., Carneiro, D. A., Borges, G. A., Chalegre, H. S., FJ, J., Silva, J. M. J. S., Silva, L. A., et al. (2013). Aprendizagem de programação através de ambientes lúdicos em um curso de engenharia de computação: Uma primeira incursão. In *Anais do XXXIII Congresso da SBC. XXI Workshop sobre Educação em Computação*, pp. 749–758.
- [Braga 2005] Braga, L. P. V. B. (2005). *Introdução à Mineração de Dados-2a edição: Edição ampliada e revisada*. Editora E-papers.
- [Brasil 2016] Brasil (2016). Censo da educação superior 2015: notas estatísticas. Disponível em: http://download.inep.gov.br/educacao_superior/censo_superior/documentos/2015/Notas_Estatisticas_Censo_Superior_2015.pdf. 18 jul. 2017.
- [Camilo e Silva 2009] Camilo, C. O. e Silva, J. C. d. (2009). Mineração de dados: Conceitos, tarefas, métodos e ferramentas. *Universidade Federal de Goiás (UFG)*, pp. 1–29.
- [Campos et al. 2012] Campos, L. C., Marques, E., e de Barros Filho, E. M. (2012). A relação entre as habilidades tecnológicas e o desempenho do aluno em cursos de educação a distância. *XV Simpósio de Administração da Produção e Operações Internacionais*.

- [Castells 2005] Castells, M. (2005). A sociedade em rede: do conhecimento à política. *A sociedade em rede. Do conhecimento à ação política. Debates–Presidência da República. Lisboa: Imprensa Nacional*, pp. 17–30.
- [Castro e Ferrari 2016] Castro, L. N. d. e Ferrari, D. G. (2016). Introdução à mineração de dados: conceitos básicos, algoritmos e aplicações. *São Paulo: Saraiva*.
- [Cerqueira et al. 2000] Cerqueira, T. C. S. et al. (2000). Estilos de aprendizagem em universitários. pp. 155.
- [Costa et al. 2014] Costa, A. R., Araújo, A. M. D. d. C., e Almeida, L. S. (2014). Envolvimento acadêmico de estudantes de engenharia: Contributos para a validação interna e externa de uma escala de avaliação. *Revista e-psi: revista eletrônica de psicologia, educação e saúde*, 1:142–155.
- [Costa et al. 2013] Costa, E., Baker, R. S., Amorim, L., Magalhães, J., e Marinho, T. (2013). Mineração de dados educacionais: conceitos, técnicas, ferramentas e aplicações. *Jornada de Atualização em Informática na Educação*, 1(1):1–29.
- [da Silva 2015] da Silva, H. G. (2015). Fatores determinantes do desempenho acadêmico no ensino superior: uma abordagem por meio do estado da arte. *VIII Encontro de Pesquisa em Educação*, pp. 15.
- [Dancey e Reidy 2018] Dancey, C. e Reidy, J. (2018). *Estatística Sem Matemática para Psicologia-7*. Penso Editora, 7th edition.
- [de Cássia Martinelli e Genari 2009] de Cássia Martinelli, S. e Genari, C. H. M. (2009). Relações entre desempenho escolar e orientações motivacionais. *Estudos de psicologia*, 14(1):13–20.
- [de Leon e Menezes-Filho 2001] de Leon, F. L. L. e Menezes-Filho, N. A. (2001). Reprovação, avanço e evasão escolar no brasil. *Cultura*, 1998(1999):36.
- [de Paula Franco 2008] de Paula Franco, A. (2008). Ensino superior no brasil: cenário, avanços e contradições. *Jornal de Políticas Educacionais*, 2(4):53–63.
- [de Souza e Coelho 2018] de Souza, E. A. e Coelho, P. M. F. (2018). Alunos querem o digital. professores, nem tanto. *Revista EducaOnline*, 12(1):67–83.
- [Dias 2002] Dias, M. M. (2002). Parâmetros na escolha de técnicas e ferramentas de mineração de dados. *Acta Scientiarum*, 24(6):1715–1725.
- [Diehl e Tatim 2004] Diehl, A. A. e Tatim, D. C. (2004). Pesquisa em ciências sociais aplicadas: métodos e técnicas (2004). pp. 58–73. São Paulo: Prentice Hall.
- [Dourado et al. 2004] Dourado, L. F., Catani, A. M., e Oliveira, J. F. (2004). Políticas públicas e reformas da educação superior no brasil: impasses e perspectivas. *Pro-Posições*, 15(3):91–115.
- [Fagundes et al. 2014] Fagundes, C. V., Luce, M. B., e Rodriguez Espinar, S. (2014). O desempenho acadêmico como indicador de qualidade da transição ensino médio-

- educação superior. *Ensaio (Rio de Janeiro (1993): avaliação e políticas públicas em educação. Vol. 22, n. 84 (jul./set. 2014), p. 635-670.*
- [Fayyad et al. 1996] Fayyad, U., Piatetsky-Shapiro, G., e Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3):37.
- [Fior e Mercuri 2018] Fior, C. A. e Mercuri, E. (2018). Envolvimento acadêmico no ensino superior e características do estudante. *Revista Brasileira de Orientação Profissional*, 19(1):85–95.
- [Fior et al. 2013] Fior, C. A., Mercuri, E., e da Silva, D. (2013). Evidências de validade da escala de envolvimento acadêmico para universitários. *Avaliação Psicológica: Interamerican Journal of Psychological Assessment*, 12(1):81–89.
- [Freitas e Hossne 2009] Freitas, C. B. D. e Hossne, W. S. (2009). O papel dos comitês de ética em pesquisa na proteção do ser humano. *Revista Bioética*, 10(2).
- [Gil 2002] Gil, A. C. (2002). Como classificar as pesquisas. volume 4, pp. 44–45. Atlas São Paulo.
- [Gouveia et al. 2010] Gouveia, V. V., Sousa, D. M. F. d., Fonseca, P. N. d., Gouveia, R. S., Gomes, A., Araújo, R. d. C. R., et al. (2010). Valores, metas de realização e desempenho acadêmico: proposta de modelo explicativo. *Revista Semestral da Associação Brasileira de Psicologia Escolar e Educacional*, 14(2):324–331.
- [Han et al. 2011] Han, J., Pei, J., e Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.
- [Hand 2007] Hand, D. J. (2007). Principles of data mining. *Drug safety*, 30(7):621–622.
- [Hand et al. 2001] Hand, D. J., Mannila, H., e Smyth, P. (2001). *Principles of data mining (adaptive computation and machine learning)*.
- [Hoed 2016] Hoed, R. M. (2016). Análise da evasão em cursos superiores: o caso da evasão em cursos superiores da área de computação. *Brasília, DF: Universidade de Brasília*, pp. 188.
- [Júnior et al. 2017] Júnior, G. C., Nascimento, R., Alves, G., e Gouveia, R. (2017). Identificando correlações e outliers entre bases de dados educacionais. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 6, pp. 694.
- [Kirschner e Karpinski 2010] Kirschner, P. A. e Karpinski, A. C. (2010). Facebook® and academic performance. *Computers in human behavior*, 26(6):1237–1245.
- [Krishnaiah et al. 2014] Krishnaiah, V., Narsimha, G., e Chandra, N. S. (2014). Survey of classification techniques in data mining. *International Journal of Computer Sciences and Engineering*, 2(9):65–74.

- [Kuh 2005] Kuh, G. D. (2005). Student engagement in the first year of college. *Challenging and supporting the first-year student: A handbook for improving the first year of college*, pp. 86–107.
- [Kuh 2009] Kuh, G. D. (2009). What student affairs professionals need to know about student engagement. *Journal of college student development*, 50(6):683–706.
- [Lemos e Lévy 2010] Lemos, A. e Lévy, P. (2010). O futuro da internet: em direção a uma ciberdemocracia planetária. *São Paulo: Paulus*, 13.
- [Levy 2010] Levy, P. (2010). *Cibercultura*. Editora 34.
- [Lu et al. 1995] Lu, H., Setiono, R., e Liu, H. (1995). Neurorule: A connectionist approach to data mining. In *Proceedings of the 21th International Conference on Very Large Data Bases*, pp. 478–489. Morgan Kaufmann Publishers Inc.
- [Lévy 1999] Lévy, P. (1999). O futuro da internet: em direção a uma ciberdemocracia planetária. são paulo: Paulus, 2010. ... *Cibercultura*. São Paulo: Editora, 34.
- [Manhães 2015] Manhães, L. M. B. (2015). Predição do desempenho acadêmico de graduandos utilizando mineração de dados educacionais. *Doutorado em Engenharia de Sistemas e Computação Instituição de Ensino: Universidade Federal do Rio de Janeiro, Rio de Janeiro. Biblioteca Depositária: BIBLIOTECA DO CT*, pp. 157.
- [MEC 1997] MEC, M. d. E. (1997). Diplomação, retenção e evasão nos cursos de graduação em instituições de ensino superior públicas. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=27010. 06 out. 2017.
- [MEC 2014] MEC, M. d. E. (2014). Documento orientador para a superação da evasão e retenção na rede federal de educação profissional, científica e tecnológica. Disponível em: <http://r1.ufrrj.br/ctur/wp-content/uploads/2017/03/Documento-Orientador-SETEC.pdf>. 08 out. 2017.
- [MEC 2017] MEC, M. d. E. (2017). Censo da educação superior: notas estatísticas 2017. Disponível em: https://www.abmes.org.br/arquivos/documentos/censo_da_educacao_superior_2017-notas_estatisticas2.pdf. 20 out. 2018.
- [Menezes Silva 2010] Menezes Silva, N. (2010). Sucesso e insucesso na educação superior: as representações sociais dos estudantes do centro acadêmico do agreste/ufpe. Master's thesis, Universidade Federal de Pernambuco.
- [Miranda et al. 2015] Miranda, G. J., da Silva Lemos, K. C., de Oliveira, A. S., e Ferreira, M. A. (2015). Determinantes do desempenho acadêmico na área de negócios. *Revista Meta: Avaliação*, 7(20):175–209.

- [Moleta et al. 2017] Moleta, D., Ribeiro, F., e Clemente, A. (2017). Fatores determinantes para o desempenho acadêmico: Uma pesquisa com estudantes de ciências contábeis. *Revista Capital Científico-Eletrônica (RCC?)*-ISSN 2177-4153, 15(3):24–41.
- [Munhoz 2004] Munhoz, A. M. H. (2004). *Uma análise multidimensional da relação entre inteligência e desempenho acadêmico em universitários ingressantes. 2004*. PhD thesis.
- [Nogueira et al. 2013] Nogueira, D. R., da Costa, J. M., Takamatsu, R. T., e dos Reis, L. G. (2013). Fatores que impactam o desempenho acadêmico: uma análise com discentes do curso de ciências contábeis no ensino presencial. *RIC-Revista de Informação Contábil-ISSN: 1982-3967*, 7(3):51–62.
- [Pascoal et al. 2016] Pascoal, T., de Brito, D. M., Andrade, L., e do Rêgo, T. G. (2016). Evasão de estudantes universitários: diagnóstico a partir de dados acadêmicos e socioeconômicos. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 27, pp. 926.
- [Pasta 2011] Pasta, A. (2011). Aplicação da técnica de data mining na base de dados do ambiente de gestão educacional: um estudo de caso de uma instituição de ensino superior de blumenau, sc. pp. 153.
- [Pimentel et al. 2003] Pimentel, E. P., de França, V. F., e Omar, N. (2003). A identificação de grupos de aprendizes no ensino presencial utilizando técnicas de clusterização. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 1, pp. 495–504.
- [Porto e Gonçalves 2017] Porto, R. C. e Gonçalves, M. P. (2017). Motivação e envolvimento acadêmico: um estudo com estudantes universitários. *Psicologia Escolar e Educacional*, 21(3):515–522.
- [Prodanov e de Freitas 2013] Prodanov, C. C. e de Freitas, E. C. (2013). *Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico-2ª Edição*. Editora Feevale.
- [Rangel e Miranda 2016] Rangel, J. R. e Miranda, G. J. (2016). Desempenho acadêmico e o uso de redes sociais. *Sociedade, Contabilidade e Gestão*, 11(2).
- [Reis e Reis 2002] Reis, E. A. e Reis, I. A. (2002). Análise descritiva de dados. *Síntese numérica Estatística*, pp. 64.
- [Ribeiro et al. 2016] Ribeiro, J. B. P., da Silva Portes, S., da Silva, A. M., e Moreira, J. R. (2016). Acesso à educação superior: tentativas de democratização do ensino em uma quase democracia. *OUTRAS PALAVRAS*, 12(1):25–31.
- [Rodrigues et al. 2017] Rodrigues, R., Gomes, A. S., e Adeodato, P. (2017). Uma abordagem de mineração de dados educacionais para previsão de desempenho a partir de padrões comportamentais de autorregulação da aprendizagem. In *Anais*

- dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 6, pp. 13.
- [Sachin e Vijay 2012] Sachin, R. B. e Vijay, M. S. (2012). A survey and future vision of data mining in educational field. In *Advanced Computing & Communication Technologies (ACCT), 2012 Second International Conference on*, pp. 96–100. IEEE.
- [Sales et al. 2014] Sales, A., Calado, B., Silva, D. R., de Oliveira Mattos, G., e Moreira, J. A. (2014). Dificuldades para o ingresso e permanência na ciência e engenharia da computação: um olhar feminino. *18º REDOR, UFRPE. Recife, PE*, pp. 3468–3482.
- [Siegel e Castellan Jr 1975] Siegel, S. e Castellan Jr, N. J. (1975). *Estatística não-paramétrica para ciências do comportamento*. Artmed Editora, 2th edition.
- [Silva et al. 2012] Silva, D. d., Ranciaro Neto, A., Anjos, L. d., e Miranda, L. (2012). Redes sociais e o desempenho acadêmico: um estudo com alunos de contabilidade. *Anais do Encontro Nacional dos Cursos de Pós-Graduação em Administração*.
- [Silva Filho et al. 2007] Silva Filho, R. L. L., Motejunas, P. R., Hipólito, O., e Lobo, M. B. C. M. (2007). A evasão no ensino superior brasileiro. *Cadernos de pesquisa*, 37(132):641–659.
- [Smith 2019] Smith, C. (2019). 65 amazing whatsapp statistics and facts (2019) | by the numbers. Disponível em: <https://expandedramblings.com/index.php/whatsapp-statistics/>. 04 abr. 2019.
- [Souza e Schneider 2014] Souza, A. A. N. e Schneider, H. N. (2014). Potencialidades do uso de sites de redes sociais no processo de ensino e aprendizagem. *International Journal of Knowledge Engineering and Management (IJKEM)*, 3(6):181–196.
- [Tavares et al. 2016] Tavares, M. R. N., Wolff, D. L., e da Silva, C. A. (2016). Forma de ingresso x situação dos alunos no curso superior de tecnologia em sistemas para internet do ifrs. # *Tear: Revista de Educação, Ciência e Tecnologia*, 5(2).
- [Wang et al. 2014] Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., e Campbell, A. T. (2014). Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 3–14. ACM.
- [Witten et al. 2016] Witten, I. H., Frank, E., Hall, M. A., e Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- [Wolf-Wendel et al. 2009] Wolf-Wendel, L., Ward, K., e Kinzie, J. (2009). A tangled web of terms: The overlap and unique contribution of involvement, engagement, and integration to understanding college student success. *Journal of College Student Development*, 50(4):407–428.

Apêndice A

Termo de Consentimento Livre e Esclarecido - TCLE (Docente)

Título do projeto: Análise da relação entre os fatores internos e externos e o desempenho acadêmico dos estudantes do curso de Engenharia da Computação da UEFS.

Mestranda: Naan Silva Cardoso

Orientador: João Batista da Rocha Junior.

Coorientadora: Claudia Pinto Pereira

Pesquisadores colaboradores: Matheus Moura Batista, Gabriel Antônio Pereira dos Santos Carneiro.

Prezado (a) professor:

Venho através deste, convidá-lo a participar do presente projeto que tem como objetivo analisar como os fatores internos e externos estão correlacionados com o desempenho acadêmico (aprovação/reprovação) do estudante. Para isso, informações como (latitude, longitude, luminosidade, data da coleta e nome do aplicativo utilizado) serão coletadas a partir do uso de um aplicativo instalado no celular do estudante, a coleta será feita dentro e fora da sala aula. Nas observações feitas nas aulas, serão coletados dados sobre (horário de chegada e saída da sala, horário de saída e retorno da sala durante a aula, o local que o aluno sentou na sala de aula, o número de questionamentos realizados, uso de recursos na aula pelo professor) durante um semestre. O professor não fará uso do aplicativo, os dados coletados serão de todas as aulas observadas no primeiro semestre de 2018. Através desse estudo, será possível buscar relações entre os dados e o desempenho acadêmico do estudante dentro e fora da sala de aula. Os dados coletados na pesquisa não serão retidos, após finalização do projeto os mesmos serão descartados. Sendo assim, os dados não serão utilizados em pesquisas futuras. Quanto aos benefícios inerentes a pesquisa, estão aqueles relacionados ao processo de ensino-aprendizagem de maneira

geral, tais como a participação dos estudantes em sala de aula, a compreensão de que os fatores podem ter relação com seu desempenho, inferir sobre dados "não aparentes" em grandes volumes de dados, através da mineração de dados, diante dos aspectos observados e dos dados coletados, refletir sobre possíveis soluções que contribuam para o processo de ensino-aprendizagem e para o desempenho dos alunos. Quanto aos riscos, existe o possível desconforto do participante em ser observado e em utilizar a ferramenta. Neste caso, a qualquer momento ele pode desistir da pesquisa sem prejuízo para ele e/ou para a disciplina observada. Além disso, os pesquisadores estarão atentos para qualquer desconforto causado, tentando saná-los. Se os mesmos permanecerem, a pesquisa poderá ser interrompida imediatamente, sem qualquer tipo de penalidade. Ademais, o anonimato será mantido, respeitando a integridade intelectual, social e cultural de cada um dos participantes. Você terá direito ao ressarcimento de qualquer gasto que tenha para a realização da pesquisa, e direito à indenização, através de assistência, caso venha a sofrer algum dano à dimensão física, psíquica, moral, intelectual, social, cultural ou espiritual, em qualquer pesquisa e dela decorrente. Caso haja danos decorrentes dos riscos previstos, o pesquisador assumirá a responsabilidade pelos mesmos. Não haverá remuneração ou qualquer custo com a participação na pesquisa. A escolha em participar desta pesquisa é livre e, se permitida, pedimos autorização de divulgação dos dados analisados em eventos científicos ou publicação em revistas, congressos e outros meios, lembrando que será mantido sigilo absoluto a respeito de seus dados pessoais. Caso haja qualquer dúvida antes, durante ou depois da realização da pesquisa, você poderá saná-la através do contato do pesquisador responsável, indicado abaixo. Em qualquer dúvida, no ponto de vista ético, o sujeito poderá consultar o Comitê de Ética em Pesquisa da Universidade Estadual de Feira de Santana através do telefone (75) 3161-8067 ou e-mail cep@uefs.br. Caso aceite participar desta pesquisa, indique o seu nome completo e assine as duas vias deste termo. Uma cópia será sua e a outra, da pesquisadora.

Feira de Santana, Ba — de ————— de 2018.

Assinatura do participante

Assinatura do pesquisador responsável
João Batista da Rocha Junior

Assinatura do responsável legal (se for o caso)

Assinatura do pesquisador colaborador

Apêndice B

Termo de Consentimento Livre e Esclarecido - TCLE (Discente)

Título do projeto: Análise da relação entre os fatores internos e externos e o desempenho acadêmico dos estudantes do curso de Engenharia da Computação da UEFS.

Mestranda: Naan Silva Cardoso

Orientador: João Batista da Rocha Junior.

Coorientadora: Claudia Pinto Pereira

Pesquisadores colaboradores: Matheus Moura Batista, Gabriel Antônio Pereira dos Santos Carneiro.

Prezado (a) estudante:

Venho através deste, convidá-lo a participar do presente projeto que tem como objetivo analisar como os fatores internos e externos estão correlacionados com o desempenho acadêmico (aprovação/reprovação) do estudante. Para isso, informações como (latitude, longitude, luminosidade, data da coleta e nome de qualquer aplicativo utilizado) serão coletadas a partir do uso de um aplicativo instalado no celular do estudante, a coleta será feita dentro e fora da sala aula. Nas observações feitas nas aulas, serão coletados dados sobre (horário de chegada e saída da sala, horário de saída e retorno da sala durante a aula, o local que o aluno sentou na sala de aula, o número de questionamentos realizados, uso de recursos na aula pelo professor) durante um semestre. Através desse estudo, será possível buscar relações entre os dados e o desempenho acadêmico do estudante dentro e fora da sala de aula. Os dados coletados na pesquisa não serão retidos, após finalização do projeto os mesmos serão descartados. Sendo assim, os dados não serão utilizados em pesquisas futuras. Quanto aos benefícios inerentes a pesquisa, estão aqueles relacionados ao processo de ensino aprendizagem de maneira geral, tais como a participação dos estudantes em sala de aula, a compreensão de que os fatores podem

ter relação com seu desempenho, inferir sobre dados "não aparentes" em grandes volumes de dados, através da mineração de dados, diante dos aspectos observados e dos dados coletados, refletir sobre possíveis soluções que contribuam para o processo de ensino aprendizagem e para o desempenho dos alunos. Quanto aos riscos, existe o possível desconforto do participante em ser observado e em utilizar a ferramenta. Neste caso, a qualquer momento ele pode desistir da pesquisa sem prejuízo para ele e/ou para a disciplina observada. Além disso, os pesquisadores estarão atentos para qualquer desconforto causado, tentando saná-los. Se os mesmos permanecerem, a pesquisa poderá ser interrompida imediatamente, sem qualquer tipo de penalidade. Ademais, o anonimato será mantido, respeitando a integridade intelectual, social e cultural de cada um dos participantes. Você terá direito ao ressarcimento de qualquer gasto que tenha para a realização da pesquisa, e também direito à indenização, através de assistência, caso venha a sofrer algum dano à dimensão física, psíquica, moral, intelectual, social, cultural ou espiritual, em qualquer pesquisa e dela decorrente. Caso haja danos decorrentes dos riscos previstos, o pesquisador assumirá a responsabilidade pelos mesmos. Se você não quiser participar da pesquisa, não haverá penalidade na disciplina. Não haverá remuneração ou qualquer custo com a participação na pesquisa. A escolha em participar desta pesquisa é livre e, se permitida, pedimos autorização de divulgação dos dados analisados em eventos científicos ou publicação em revistas, congressos e outros meios, lembrando que será mantido sigilo absoluto a respeito de seus dados pessoais. Caso haja qualquer dúvida antes, durante ou depois da realização da pesquisa, você poderá saná-la através do contato do pesquisador responsável, indicado abaixo. Em qualquer dúvida, no ponto de vista ético, o sujeito poderá consultar o Comitê de Ética em Pesquisa da Universidade Estadual de Feira de Santana através do telefone (75) 3161-8067 ou e-mail cep@uefs.br. Caso aceite participar desta pesquisa, indique o seu nome completo e assine as duas vias deste termo. Uma cópia será sua e a outra, do pesquisador. Caso você seja menor de idade, não poderá participar desta pesquisa.

Feira de Santana, Ba — de ————— de 2018.

Assinatura do participante

Assinatura do pesquisador responsável
João Batista da Rocha Junior

Assinatura do responsável legal (se for o caso)

Assinatura do pesquisador colaborador