



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS: CORRENTE
CURSO: ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

GABRIEL SANTOS BORGES

**UTILIZAÇÃO DE APRENDIZEM DE MÁQUINA PARA A IDENTIFICAÇÃO DE
DEPENDÊNCIA EM APARELHOS CELULARES COM FOCO EM CASOS QUE
POSSAM CAUSAR REPROVAÇÃO E EVASÃO**

CORRENTE

2021

GABRIEL SANTOS BORGES

Utilização de aprendizagem de máquina para a identificação de dependência em aparelhos celulares com foco em casos que possam causar reprovação e evasão

Trabalho de Conclusão de Curso apresentado como exigência parcial para obtenção do diploma do Curso de Análise e desenvolvimento de sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí. Campus Corrente.

Orientador: Prof. Ms. Felipe Gonçalves do Santos.
Coorientador : Prof. Dr. Thiago Reis da Silva
Coorientador : Psic. Ms. Aline Rocha Bezerra

CORRENTE

2021

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD

Borges, Gabriel Santos

B118u Utilização de aprendizagem de máquina para a identificação de dependência em aparelhos celulares com foco em casos que possam causar reprovação e evasão / Gabriel Santos Borges. - 2021.
24 p.: il. color.

Trabalho de Conclusão de Curso (Tecnologia em Análise e Desenvolvimento de Sistemas) - Instituto Federal do Piauí, Campus Corrente, 2021.

Orientador : Prof. Me. Felipe Gonçalves dos Santos .

Coorientador : Prof. Dr. Thiago Reis da Silva.

1. Análise e Desenvolvimento de Sistemas - Estudo e Ensino. 2. Aprendizagem de máquina . 3. Nomofobia. 4. Evasão Escolar . 5. Reprovação escolar. I. Título.

CDD - 004.21

Elaborado por Tefischer Huanderson Soares e Sousa CRB 3/1251

GABRIEL SANTOS BORGES

Utilização de aprendizagem de máquina para a identificação de dependência em aparelhos celulares com foco em casos que possam causar reprovação e evasão.

Trabalho de Conclusão de Curso apresentado como exigência parcial para obtenção do diploma do Curso de Análise e desenvolvimento de sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí. Campus Corrente.

Aprovada em: ____/____/____.

BANCA EXAMINADORA

Prof. Ms. Felipe Gonçalves do Santos (Orientador)
Instituto Federal do Piauí (IFPI)

Prof. Prof. Dr. Thiago Reis da Silva (Coorientador)
Instituto Federal do Maranhão (IFMA)

Prof. Prof. Dr. Jonathas Jivago de Almeida Cruz
Instituto Federal do Piauí (IFPI)

Prof. Prof. Dr. Karl Hansmuller Alelaf Ferreira
Instituto Federal do Piauí (IFPI)

RESUMO

Este trabalho introduz os problemas causados pelo uso excessivo dos celulares na sociedade moderna, e mostra que eles podem causar problemas escolares aos alunos que o usam de forma demasiada. Para isto, foi feita uma integração da tecnologia como uso de uma área da inteligência artificial, o aprendizado de máquina, com dados de uma base de dados específica. Para a tarefa proposta foram usados os classificadores Naive Bayes, AdaBoost, SVM, Bagging e Random Forest, ao final dos testes o classificador SVM apresentou o melhor desempenho no contexto geral. Com as taxas variando de 91,22 a 80 por cento nos três cenários testados.

Palavras-chave: Aprendizado de máquina, nomofobia, evasão escolar.

ABSTRACT

This work introduces the problems caused by the excessive use of cell phones in modern society and shows that they can cause school problems to students who use it too much. For this, we integrated technology with the use of an artificial intelligence area, machine learning, with data from a specific database. For the proposed task, the Naive Bayes, AdaBoost, SVM, Bagging and Random Forest classifiers were used. At the end of the tests, the SVM classifier presented the best performance in the general context. With rates ranging from 91.22 to 80 percent in the three scenarios tested.

Keywords: Machine learning, nomophobia, school dropout.

LISTA DE ILUSTRAÇÕES

Figura 1	Aprendizado de máquina supervisionado	15
Figura 2	Etapas da abordagem proposta	14
Gráfico 1	Dados dos alunos do primeiro ano do ensino médio.....	20
Gráfico 2	Dados dos alunos do segundo ano do ensino médio.....	20
Gráfico 3	Dados dos alunos do segundo ano do ensino médio.....	21

LISTA DE TABELAS

Tabela 1	Resultados dos primeiros anos: caso em que possuem uma possível dependência	17
Tabela 2	Resultados dos segundos anos: caso em que possuem uma possível dependência	17
Tabela 3	Resultados dos terceiros anos: caso em que possuem uma possível dependência	17
Tabela 4	Resultados dos 1º anos: casos em que não são classificados com dependência	18
Tabela 5	Resultados dos 2º anos: casos em que não são classificados com dependência	18
Tabela 6	Resultados dos terceiros anos: casos em que não são classificados com dependência	19

LISTA DE ABREVIATURAS E SIGLAS

IBGE Instituto Brasileiro de Geografia e Estatística

IFPI Instituto Federal de Educação, Ciência e Tecnologia do Piauí

SUMÁRIO

1	INTRODUÇÃO	11
2	TRABALHOS CORRELATOS	12
3	MATERIAIS E MÉTODOS	14
3.1	ABORDAGEM PROPOSTA	15
3.2	COLETA DE DADOS	15
3.3	<i>PROCESSAMENTO</i>	15
3.4	<i>CLASSIFICAÇÃO</i>	16
3.5	<i>TECNOLOGIAS UTILIZADAS</i>	16
3.6	<i>VALIDAÇÃO</i>	16
4	RESULTADOS E DISCUSSÕES	16
5	CONSIDERAÇÕES FINAIS	21
	REFERÊNCIAS	22
	APÊNDICE A – INSTRUMENTO DE COLETA DE DADOS	24

1 INTRODUÇÃO

A evolução tecnológica no mundo, e particularmente no Brasil, vem sendo importante para o desenvolvimento econômico, cultural e social. “Segundo dados do IBGE, referentes aos três últimos meses de 2018, tem-se que o percentual de brasileiros com acesso à internet aumentou no país de 2017 para 2018, passando de 69,8% para 74,7%, mas que 25,3% ainda estão sem acesso.” (TOKARNIA, 2020). Ainda segundo os dados apresentados, “tem-se que o índice de domicílios com acesso à internet também apresentou um aumento entre 2017 e 2018, passando de 74,9% para 79,1%.” (TOKARNIA, 2020). “Sendo que o percentual de pessoas de 10 anos ou mais que acessaram a internet pelo celular passou de 97% para 98,1%.” (TOKARNIA, 2020).

Neste contexto, o uso em excesso de aparelhos telefônicos pode causar problemas tanto físicos quanto emocionais, como pode se observar em Soares (2019), entre os problemas no âmbito emocional, se destacam a depressão, ansiedade, impulsividade e até a diminuição da capacidade social. Entre os problemas físicos demonstrados em Sousa e Miranda (2018), se destacam dores de cabeça, nos braços, no pescoço e as contraturas dos músculos dorsais. Além destes, é apresentado em Nina (2016), os problemas causados nos ambientes escolares, onde eles podem ser responsáveis pela queda de desempenho dos alunos que fazem uso de forma demasiada do mesmo. Essa dependência em smartphones é denominada nomofobia, que segundo Faria (2019) é um termo que descreve o medo de ficar sem contato constante com o celular.

Através destes dados, pode-se observar como o uso de *smartphones* podem ser prejudiciais aos alunos, pois o mesmo causa diversas distrações a quem os possuem, gerando uma certa dependência, ocasionando assim uma possível queda no desempenho acadêmico dos estudantes. Este artigo busca através do Aprendizado de Máquina (*Machine Learning*), realizar o diagnóstico da dependência por aparelhos *mobile* através de um conjunto de pesquisas desenvolvidas junto a psicóloga no Instituto Federal de Educação, Ciência e Tecnologia do Piauí – Campus Corrente.

“Aprendizado de Máquina (*Machine Learning*) é uma área de IA cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado bem como a construção de sistemas capazes de adquirir conhecimento de forma automática.” (MONARD & BARANAUSKAS, 2003, P. 39). A mesma também pode ser definida como “uma subárea da Inteligência Artificial (IA), que estuda o desenvolvimento de métodos capazes de extrair informações e conhecimento a partir de amostras de dados”(TAKAKURA, 2018, P. 2).

O uso de técnicas de machine learning neste contexto tem como principal objetivo identificar jovens que possam ter um certo grau de dependência em aparelhos celulares. Para reportar o trabalho realizado, as demais seções deste artigo encontram-se organizadas da seguinte forma. Na Seção 2 são reportados os trabalhos correlatos a esta pesquisa. Já a Seção 3 aborda os materiais e métodos, a Seção seguinte os resultados e discussões e, por fim, na Seção 5, são apresentadas as considerações finais.

2 TRABALHOS CORRELATOS

No estudo de Elden *et. al.*, (2013), foi proposto o desenvolvimento de um novo algoritmo chamado Ada-GA, que utiliza AdaBoost *Ensemble* com um algoritmo genético para construir um modelo preditivo de classificação de desempenho de estudantes do ensino superior. O Ada-GA é um algoritmo de *boosting* que objetiva otimizar a quantidade de classificadores fracos e seus pesos usando algoritmo genético, com o intuito de melhorar o desempenho do comitê de classificadores. Após vários experimentos utilizando algoritmos como J48, *Naive Update*, KMeans para compor o comitê com o AdaBoost, os autores apresentaram uma tabela de comparação de resultados, onde o AdaBoost teve um percentual de precisão preditiva de 81,85% e o modelo proposto apresentou 82,07% de precisão.

Na pesquisa de Rosales (2017), propôs um projeto no qual foi desenvolvido um *framework* capaz de determinar o grau da severidade de experiências relatadas do *bullying* com o limite de até 140 caracteres. Em um primeiro momento eram classificados com *Support Vector Machine* (SVM) um classificador binário supervisionado, sendo posteriormente atribuído o desenvolvimento do sistema de Lógica Fuzzy.

Para melhoria dos SVM, eram rotulados múltiplos textos para aumentara acurácia e precisão das classificações. Com a precisão de 95,42% de acerto na predição, considerado que se precisava de melhora nas classificações do SVM.

No estudo de Manhães *et. al.*, (2011), foram utilizadas técnicas de mineração de dados para extrair informações de evasão/retenção da base de dados do SIG@ de estudantes dos cursos de graduação da Universidade Federal do Rio de Janeiro, identificando três classes distintas de alunos. A pesquisa avaliou um conjunto de seis algoritmos de mineração de dados do software Weka: Árvore de decisão (J48 e *Simple Chart*), SVM, Estatístico (*Naive Bayes*), Redes neurais artificiais (MLP), *Ensemble* (AdaBoost), os quais apresentaram uma acurácia entre 70% a 86%. Ao final eles adotaram o *Naive Bayes* que, apesar de não ter sido o melhor classificador, apresentou um rendimento global que atendia aos objetivos da pesquisa.

No trabalho de Brandão *et. al.*, (2017) é feito uma abordagem de aprendizagem de máquinas com foco em estudantes com retenção universitária, um problema que preocupa os diversos atores envolvidos com educação do ensino superior. Dessa forma, foi desenvolvido um modelo de predição utilizando individualmente os classificadores J48, SVM, RBF, Naive Bayes e MLP, para posteriormente, formar um comitê desses classificadores, testes de hipótese e validação usando a ferramenta Weka e, após a escolha da medida de desempenho Recall, a qual melhor se adequa ao problema em questão, pode-se inferir que o modelo de árvore de decisão gerada pelo algoritmo J48 generalizou melhor que os outros classificadores, mesmo quando combinados em Ensemble com o método Bagging.

Nos trabalhos referenciados nota-se que os autores buscaram com técnicas de aprendizagem de máquina diminuir o número de evasões escolares, isto é feito através de diferentes métodos. Nesta pesquisa é abordado como o uso constante de celulares pode vir a ser responsável pelo problema de evasão e reprovação escolar, e junto as técnicas de *Machine Learning* busca-se encontrar padrões para ocasionar essas situações e através destes resultados apresentar aos responsáveis, para que com posse dos dados apresentados estes possam vir a auxiliar na diminuição do impacto deste problema no ambiente escolar.

3 MATERIAIS E MÉTODOS

“A aprendizagem de máquina é um sub-campo da inteligência artificial dedicado ao desenvolvimento de algoritmos e técnicas que permitam ao computador aprender, isto é, que permitam ao computador aperfeiçoar seu desempenho em alguma tarefa” (AMORIM, 2008, P. 1).

A tecnologia de aprendizagem de máquinas pode ser categorizada como: supervisionada ou não supervisionada. Takahashi et. al.(2015) define as mesmas da seguinte forma: o algoritmo supervisionado recebe exemplos de entradas e saídas e a partir disso “aprende” uma regra que mapeia as entradas e saídas, como pode ser observado na Figura 1. Já no não supervisionado, não são fornecidos exemplos de saída, somente de entrada, então é o próprio algoritmo que define os agrupamentos e tipos de saídas a serem devolvidos. Existem diversos tipos de algoritmos de aprendizagem de máquina, cada um com uma metodologia e finalidades diferentes, como exemplo tem-se, o *Naïve Bayes*, *AdaBoost*, dentre outros.

Dentre as várias abordagens existentes de aprendizado de máquina, a preditiva também conhecida como supervisionada foi a escolhida para este estudo, pois esta se encaixava mais no propósito desta pesquisa, pois para determinar o melhor resultado fez se necessário que os classificadores treinassem com exemplos de dados de saída. Para que os mesmos tivessem certo embasamento na etapa de classificação.

Figura 1: Aprendizado de máquina supervisionado.

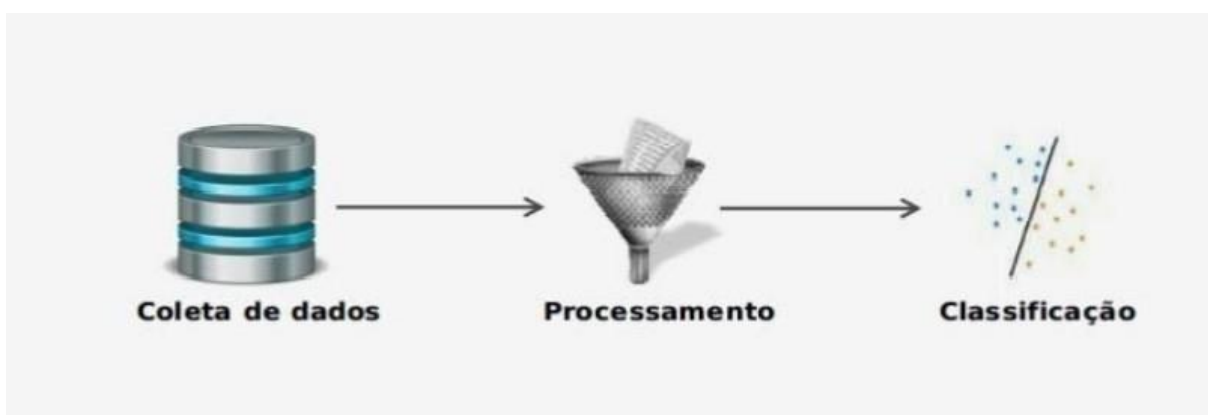


Fonte: O autor (2019).

3.1 Abordagem Proposta

A abordagem proposta neste trabalho é ilustrada na Figura 2. Para atingir o objetivo do estudo a abordagem inclui as seguintes etapas: coleta de dados, processamento e classificação.

Figura 2: Etapas da abordagem proposta.



Fonte: De Paula et. al. (2017)

3.2 Coleta de dados

Os dados usados nesta pesquisa foram recolhidos em 11 turmas do ensino médio integrado ao técnico do IFPI – Campus Corrente, através de um questionário¹ elaborado conjuntamente com a psicóloga da instituição contendo 12 questões relacionadas ao tema da pesquisa. Ao final foi obtido um total de 334 alunos, sendo 171 do sexo masculino e 163 do sexo feminino. Destes, foi determinado que 50% seria para treino e os outros 50% seriam utilizados para testes.

3.3 Processamento

Os dados usados foram processados da seguinte forma, no primeiro momento foram eliminados dados que não seriam necessários para a pesquisa determinar uma possível dependência, como exemplo tem-se a idade do aluno em, depois dessa limpeza os dados foram colocados em um arquivo no formato *Comma-separated values* (.csv) para que depois o mesmo fossem submetidos aos classificadores.

¹ Questionario de avaliação - <https://docs.google.com/forms/d/1ZJeHaK5j0iCXeLidR61Wuh766UpWdN8FdZ1cbhRGydc/edit>

3.4 Classificação

Nesta etapa os dados foram submetidos aos algoritmos de aprendizagem de máquina. Segundo De Paula *et. al.*, (2017), uma característica importante de técnicas supervisionadas é que elas apresentam fundamentalmente o mesmo comportamento: dado um conjunto de dados de treinamento, composto por instâncias que são formadas por vários atributos previamente rotulados (polaridades) nas classes de interesse, o modelo aprende características (base de treino) de cada classe e poderá ser usado para classificar outras amostras (base de teste). Os classificadores adotados foram o *Multinomial Naive Bayes*, *AdaBoostClassifier*, *Bagging*, *Máquinas de Vetores Suporte (SVM)* e o *Random Forest*.

3.5 Tecnologias utilizadas

Todas as implementações foram desenvolvidas em *Python 3*. Os algoritmos implementados fazem parte de uma biblioteca *open source* de aprendizado de máquina do *Python* chamada *Scikit-Learn*².

3.6 Validação

Para seleção do melhor conjunto de parâmetros do algoritmo de aprendizado de máquina foram utilizadas as métricas de: “precisão, revocação e o F-Measure.” (SILVA, 2018). Nas subseções a seguir discutiremos sobre os critérios citados e sua aplicação, utilizando as seguintes abreviações: TP que são os verdadeiros positivos, FN que representa os falsos negativos e FP que representa os falsos positivos.

4 Resultados e Discussões

Após analisar os dados coletados pela pesquisa nos algoritmos de aprendizagem de máquina pré-selecionados, foram alcançados os resultados apresentados nas Tabelas 1, 2, 3, 4, 5 e 6 respectivamente. Nas Tabelas 1, 2 e 3, estão dispostos os dados em que os algoritmos tinham como função avaliar se aquele indivíduo em questão possui uma certa dependência em aparelhos *mobile*. Na Tabela 4, 5 e 6, estão representados os dados quando os algoritmos tinham como objetivo classificar se a pessoa em questão não apresentava dependência em dispositivos *mobile*.

² *Scikit-Learn* - <https://scikit-learn.org/stable/>

Tabela 1 – Resultados dos primeiros anos: caso em que possuem uma possível dependência.

	Precisão	Revocação	F-Measure	Taxa de Acerto
<i>Naive Bayes</i>	0,75	0,115	0,199	56,12
<i>AdaBoost</i>	0,875	0,777	0,823	84,12
<i>SVM</i>	1,0	0,814	0,897	91,22
<i>Bagging</i>	0,777	0,518	0,621	78,94
<i>Random Forest</i>	0,894	0,629	0,738	85,96

Tabela 2 – Resultados dos segundos anos: caso em que possuem uma possível dependência.

	Precisão	Revocação	F-Measure	Taxa de Acerto
<i>Naive Bayes</i>	0,9	0,409	0,562	53,3
<i>AdaBoost</i>	0,894	0,772	0,828	76,6
<i>SVM</i>	1,0	0,727	0,841	80,0
<i>Bagging</i>	0,937	0,681	0,788	73,3
<i>Random Forest</i>	1,0	0,590	0,742	63,3

Tabela 3 – Resultados dos terceiros anos: caso em que possuem uma possível dependência.

	Precisão	Revocação	F-Measure	Taxa de Acerto
<i>Naive Bayes</i>	0,6	0,6	0,6	60,0
<i>AdaBoost</i>	0,666	0,6	0,628	53,3
<i>SVM</i>	0,875	0,7	0,777	80,0
<i>Bagging</i>	0,666	0,6	0,628	53,3
<i>Random Forest</i>	0,75	0,6	0,666	80,0

Na Tabela 1, estão dispostos os dados dos alunos das turmas de primeiro ano do ensino médio, ao analisar os resultados obtidos pode-se observar que o algoritmo SVM, foi o que apresentou o melhor resultado no contexto geral, com uma taxa de acerto e *F-Measure* de 91,22% e 0,897 respectivamente. Já nos testes realizados com as turmas de segundo e terceiro ano, como pode ser observado nas tabelas 2 e 3, o SVM foi novamente o que apresentou os melhores resultados.

Com uma taxa de acerto e *F-Measure* de 80,0% e 0,841 respectivamente nos alunos de segundo ano, mas na classificação dos alunos dos terceiros anos o *Random Forest* apresentou resultados semelhantes ao SVM, pois ambos obtiveram uma taxa de acerto de 80%, nos dois casos, com o SVM apresentado um melhor *F-Measure*. Em contrapartida, os piores resultados vieram do algoritmo *Naive Bayes* nas turmas de primeiro e segundo anos com as seguintes taxa de acertos 56,12% e 53,33%, já nos terceiros anos os piores resultados ficaram com os classificadores *AdaBoost* e *Bagging* com uma percentagem de acerto 53,33%.

Tabela 4 – Resultados dos 1º anos: casos em que não são classificados com dependência.

	Precisão	Revocação	<i>F-Measure</i>	Taxa de Acerto
<i>Naive Bayes</i>	0,547	0,966	0,698	56,12
<i>AdaBoost</i>	0,818	0,9	0,857	84,21
<i>SVM</i>	0,685	0,774	0,726	91,22
<i>Bagging</i>	0,666	0,866	0,752	78,94
<i>Random Forest</i>	0,675	0,833	0,745	85,96

Tabela 5 – Resultados dos 2º anos: casos em que não são classificados com dependência.

	Precisão	Revocação	<i>F-Measure</i>	Taxa de Acerto
<i>Naive Bayes</i>	0,35	0,318	0,333	53,3
<i>AdaBoost</i>	0,454	0,625	0,525	76,6
<i>SVM</i>	0,571	1,0	0,726	80,0
<i>Bagging</i>	0,500	0,538	0,518	73,3
<i>Random Forest</i>	0,470	1,0	0,639	63,3

Tabela 6 – Resultados dos terceiros anos: casos em que não são classificados com dependência.

	Precisão	Revocação	<i>F-Measure</i>	Taxa de Acerto
<i>Naive Bayes</i>	0,6	0,6	0,6	60,0
<i>AdaBoost</i>	0,714	0,4	0,363	53,3
<i>SVM</i>	0,714	1,0	0,833	80,0
<i>Bagging</i>	0,333	0,4	0,336	53,3
<i>Random Forest</i>	0,428	0,6	0,499	80,0

Na Tabela 4, 5 e 6, estão representados os dados dos alunos das turmas de primeiro, segundo e terceiro ano do ensino médio respectivamente, ao averiguar as informações obtidas pode-se constatar que algoritmo que apresentou o melhor desempenho foi o SVM, com uma taxa de acerto e *F-Measure* de superior aos demais praticamente todos os cenários, mas vale ressaltar que o classificador *Random Forest* apresentou uma taxa de acerto igual ao SVM na classificação dos alunos dos terceiros anos, como pode ser observado na tabela 6, mas na métrica *deF-Measure* pode-se notar que o SVM obteve um melhor desempenho no geral.

O pior desempenho ficou com os algoritmos *Naive Bayes* novamente nas turmas de primeiro e segundo ano. Nos testes realizados com as turmas de terceiros anos, os algoritmos *AdaBoost* e *Bagging*, foram os que apresentaram os piores resultados. Foi obtido após a análise de todos os resultados, que o classificador SVM é o mais apto para a tarefa. Em contraparte os piores resultados vieram pelo lado do algoritmo *Naive Bayes* com taxas baixas em relação aos outros classificadores.

Um fato que pode ser notado após a análise dos resultados foi que os algoritmos apresentaram uma queda de desempenho quando as turmas analisadas eram dos segundo e terceiro anos, isso se deve tanto pelo número de dados para treino serem mais escassos. Outra análise realizada nesta pesquisa é a porcentagem de homens e mulheres que foram avaliadas com uma possível dependência por dispositivos *mobile* e a porcentagem dos que foram avaliados na situação oposta, ou seja, os que não possuem dependência. Os dados foram divididos de acordo com as turmas de primeiro, segundo e terceiro anos, os dados estão expostos nos Gráficos 1, 2 e 3.

Para a realização da classificação do aluno nos perfis abordados na pesquisa, foram levados em conta os padrões que haviam sido discutidos antes da aplicação do questionário com a psicóloga da instituição. Os padrões para a classificação da possível dependência foram: das 12 questões presentes no questionário, ficou definido que se o aluno respondesse, sim em 5 questões, este já se encaixaria no grupo de possível dependência a dispositivos *mobile*.

Gráfico 1 – Dados dos alunos do primeiro ano do ensino médio.

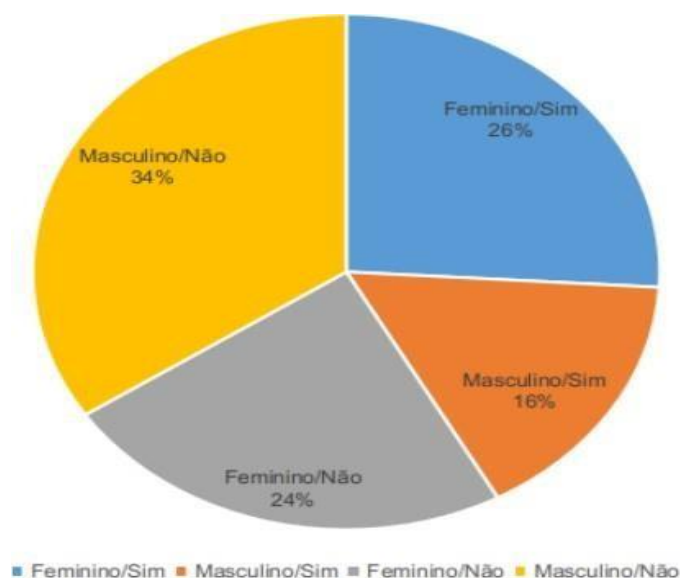


Gráfico 2 – Dados dos alunos do segundo ano do ensino médio.

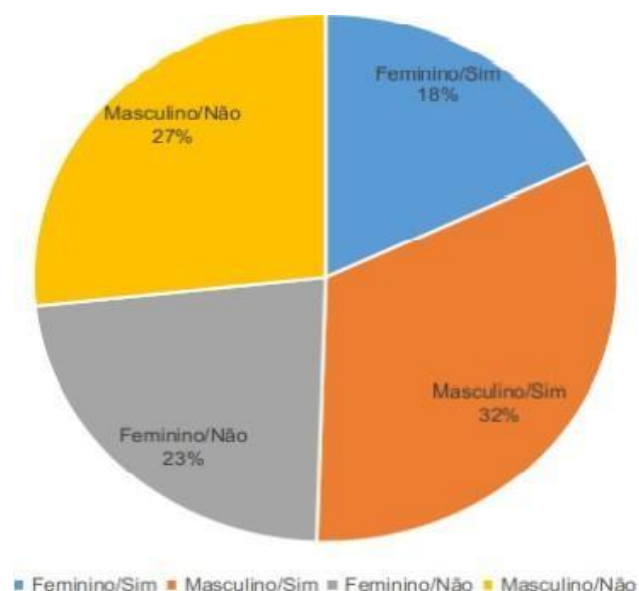
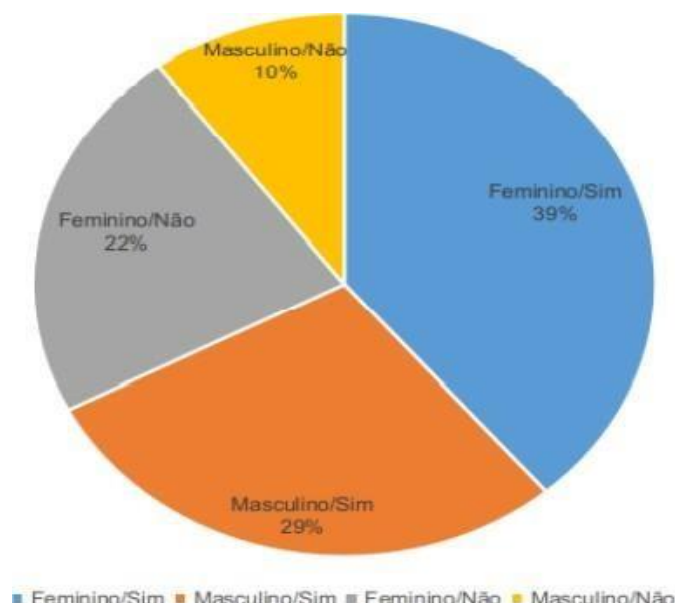


Gráfico 3 – Dados dos alunos do terceiro ano do ensino médio.



Nos gráficos 1 e 2, onde estão sendo representados os dados dos alunos de primeiros e segundos anos, o total de pessoas que foram avaliadas como dependentes do *smartphone* e às quais não foram avaliadas como dependentes foi relativamente igual, o fato um tanto quanto preocupante, pois boa parte dos alunos se dedicam ao celular como uma prioridade maior que o estudo, isso não apenas prejudica nos resultados das avaliações como também na adesão no mercado profissional para formação técnica.

O uso do *smartphone* é ainda mais alarmante nos terceiros anos (Gráfico 3), pois os alunos já estão próximos de serem formados na instituição e a um passo do mercado de trabalho, o uso do aparelho é ainda mais crítico, visto que 68% dos alunos destas turmas foram considerados com uma certa dependência a dispositivos *mobile*.

5 CONSIDERAÇÕES FINAIS

Após todos os testes realizados é válido concluir que o algoritmo SVM foi o que apresentou os melhores índices entre os cinco classificadores testados, e ao mesmo tempo foi o melhor em todos os cenários avaliados. Na contramão do SVM, o *Naive Bayes* foi o pior classificado em dois dos três cenários testados. No cenário restante os algoritmos *AdaBoost* e *Bagging* apresentaram os piores resultados.

Outro fator interessante é a integração da tecnologia como uma possível forma sanar um problema que ocorre no ambiente acadêmico, mas também pode impactar de forma positiva na vida da pessoa que aceita que possui uma certa sujeição a aparelhos celulares, pois o uso excessivo do mesmo causa impactos tanto físicos quanto mentais.

Como trabalhos futuros pretende-se aplicar este questionário em outras escolas da região e em novas turmas que forem adentrando no IFPI-Campus Corrente, para que possa-se melhorar ainda mais a base de dados, e com isto obter uma melhora nos níveis de precisão dos algoritmos para que os testes possam ser ainda mais confiáveis em todos os cenários possíveis.

REFERÊNCIAS

AMORIM, Maurício JV; BARONE, Dante; MANSUR, André Uebe. Técnicas de aprendizado de máquina aplicadas na previsão de evasão acadêmica. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2008. p. 666-674.

Brandão, J. O. S. , Silva, A. J., Gouveia, R. M. M. , Soares, R. G. F. Aprendizagem de máquina para predição de desempenho de estudantes de graduação na UFPE. **XIV Encontro Nacional de Inteligência Artificial e Computacional - SBC ENIAC-2017**. v. XIV, p. 205-216.

De Paula, Hildon et al. Quantificando a Importância de Emojis e Emoticons para Identificação de Polaridade em Avaliações Online. In: **Anais do XIV Simpósio Brasileiro de Sistemas Colaborativos**. SBC, 2017. p. 228-239.

Elden, Ahmed Sharaf et al. AdaBoost ensemble with simple genetic algorithm for student prediction model. **International Journal of Computer Science & Information Technology**, v. 5, n. 2, p. 73, 2013.

Faria, C. Nomofobia: O que é, Como identificar e Tratar. **TUASAÚDE**, 24 de jul. de 2019. Disponível em: <<https://www.tuasaude.com/nomofobia/#>>. Acesso em: 26/01/2021.

Manhães et. al. Previsão de Estudantes com Risco de Evasão Utilizando Técnicas de Mineração de Dados . **XXII SBIE - XVII WIE** 2011.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**, v. 1, n. 1, p. 32, 2003.

Nina. O uso do celular e as consequências negativas para o rendimento escolar. **Instituto Dom Barreto**, 04 de maio de 2016. Disponível em: <<http://dombarreto.g12.br/portal/o-uso-do-celular-e-as-consequencias-negativas-para-o-rendimento-escolar/>>. Acesso em: 16/05/2020.

Rosales, S. et. al. Framework para identificação da severidade de bullying baseado em machine learning e lógica fuzzy. Dissertação de Mestrado. Universidade Estadual de Campinas – UNICAMP. 2017.

SILVA, T. A. Como interpretar as métricas Precisão, Revocação, Acurácia e Medida-F. 2018. Disponível em: <<https://tiago.blog.br/precisao-revocacao-acuracia-e-medida-f/>>. Acesso em: 18/05/2020.

Soares, I. Uso abusivo dos smartphones está causando uma epidemia de ansiedade, depressão e dores. **Gazeta do Povo**, 04 de maio de 2019. Disponível em: <<https://www.gazetadopovo.com.br/viver-bem/saude-e-bem-estar/ansiedade-depressao-dores-smartphones-tambem-podem-causar-problemas/>>. Acesso em: 16/05/2020.

Sousa, A.; Miranda, A. Os problemas causados pelo uso excessivo de smartphones. 2018. **Artigo**. Disponível em: <http://prpi.ifce.edu.br/nl_lib/file/doc4147-Trabalho/ARTIGO%20terminado.pdf>. Acesso em: 16/05/2020.

Takahashi, Marcos M.et. al. Estudo comparativo de Algoritmos de Recomendação. **Universidade de São Paulo**, 2015.

TAKAKURA, Aline Miki et al. USO DO APRENDIZADO DE MÁQUINA NO DIAGNÓSTICO MÉDICO DE PATOLOGIAS. In: **Colloquium Exactarum**. ISSN: **2178-8332**. 2018. p. 78-89.

TOKARNIA, M. Um em cada 4 brasileiros não tem acesso à internet, mostra pesquisa. **Agência Brasil**, 29 de maio de 2020. Disponível em: <<https://agenciabrasil.ebc.com.br/economia/noticia/2020-04/um-em-cada-quatro-brasileiros-nao-tem-acesso-internet>>. Acesso em: 18/05/2020.

TOKARNIA, M. Celular é o principal meio de acesso à internet no país. **Agência Brasil**, 29 de maio de 2020. Disponível em: <<https://agenciabrasil.ebc.com.br/economia/noticia/2020-04/celular-e-o-principal-meio-de-acesso-internet-no-pais#:~:text=O%20uso%20do%20celular%20para%20acessar%20a%20internet%20cresceu%20no%20Brasil.&text=Os%20dados%20mostram%20que%2079,%C3%A0%20rede%20por%20meio%20deles>>. Acesso em: 18/05/2020.

APÊNDICE A – INSTRUMENTO DE COLETA DE DADOS

QUESTIONÁRIO

1. Já me disseram mais de uma vez que eu passo tempo demais no smartphone?
() Sim () Não
2. Eu me sinto desconfortável/ansioso/inquieto quando eu fico sem usar o smartphone durante um certo período de tempo?
() Sim () Não
3. Eu me sinto disposto a usar o smartphone mesmo quando me sinto cansado?
() Sim () Não
4. Eu uso smartphone durante mais tempo e/ou gasto mais dinheiro nele do que eu pretendia inicialmente?
() Sim () Não
5. Eu tenho tentado passar menos tempo usando o smartphone, mas não tenho conseguido?
() Sim () Não
6. Eu não consigo controlar o impulso de utilizar o smartphone?
() Sim () Não
7. Eu me sinto mais satisfeito utilizando o smartphone do que passando tempo com meus amigos?
() Sim () Não
8. Eu sinto dores ou incômodos nas costas, ou desconforto nos olhos, devido ao uso excessivo do smartphone?
() Sim () Não
9. A ideia de utilizar o smartphone vem como primeiro pensamento na minha cabeça quando acordo de manhã?
() Sim () Não
10. O uso de smartphone tem causado efeitos negativos no meu desempenho na escola ou no trabalho?
() Sim () Não
11. Minha vida seria sem graça se eu não tivesse o smartphone?
() Sim () Não
12. Navegar no smartphone tem causado prejuízos para a minha saúde física. Por exemplo, uso o smartphone quando atravesso?
() Sim () Não