

Relatório de Projeto Final da disciplina de Inteligência Artificial aplicada à Engenharia Biomédica

Deidre Imani Taylor, Laís Souza Amorim, Victor Hugo de Lima Kunst

Resumo — Este relatório tem como objetivo registrar os procedimentos realizados para análise da base de dados em torno do problema da qualidade de saúde mental de pessoas que trabalham na área de tecnologia.

I. INTRODUÇÃO

A saúde mental no trabalho é definida pela OMS como “um estado de bem-estar em que o indivíduo está ciente de suas próprias habilidades, pode enfrentar as tensões normais da vida, pode trabalhar de forma produtiva e frutífera e é capaz de contribuir com a sua comunidade”. E dentro do trabalho tecnológico, essa problemática se mostra extremamente preocupante, tendo em vista a quantidade alarmante de casos, por exemplo, de Burnout em profissionais de TI [1-3] e o fato de que saúde mental ainda é considerada tabu no ambiente de trabalho [4].

Levando em consideração essas informações, foi escolhido a base de dados “Mental Health in Tech Survey”, uma pesquisa de saúde mental na tecnologia realizada em 2014 e que procura entender o comportamento, histórico, e sinais que os profissionais podem apresentar para tentar prever e, consequentemente, ajudar a prevenir que esse problema venha ou continue a ocorrer. Diferentes fatores afetam a saúde mental. Embora os parâmetros individuais não sejam separados dos parâmetros ambientais, existem classificações abrangentes nas quais fatores individuais (por exemplo, fatores genéticos, gênero, fisiologia, crenças individuais e habilidades pessoais) e fatores ambientais (por exemplo, condições culturais, relações sociais, ambiente de vida e local de trabalho)) são considerados fatores decisivos [5]. Para isto, a base é analisada, verificando cada atributo, o que podem sugerir e sua importância. Após a análise, é aplicado uma adaptação dos atributos, em seguida uma seleção é feita utilizando algoritmos bioinspirados e, por fim, vários algoritmos são aplicados a fim de encontrar a melhor solução para o problema.

Inicialmente algumas hipóteses foram levantadas enquanto exploramos o problema e a base de dados:

- **O impacto do atributo de gênero nos resultados:** Muito provavelmente mulheres (cis) e pessoas não cisgêneras possuem maior propensão a sofrerem mentalmente no meio tecnológico, e automaticamente a buscarem ajuda psicológica, então este atributo deve ser importante;
- **A época em que foi feita a pesquisa:** A pesquisa foi realizada em 2014, ou seja, provavelmente o tabu com relação a saúde mental era ainda mais frequente, e a possibilidade de busca de ajuda talvez fosse ainda menos disseminada;
- **A influência da desproporção dos dados com relação aos estados e países:** A análise com relação a localização geográfica pode ter sido dificultada devido ao desbalanceamento dos dados, já que a maior parte dos participantes (aproximadamente 60%) eram dos Estados Unidos;
- **Os atributos que achamos inicialmente que seriam relevantes para análise:** Age, Gender, Family History, Tech Company, Wellness Program, Leave, Mental Health Consequence, mental_vs_physical.

II. METODOLOGIA

Para a realização deste trabalho, optamos pela seguinte sequência de etapas:

1. **Entender o problema e a base de dados:** Nesta etapa o dataset é explorado, a fim de tentar compreender como ele foi construído, o significado de cada um dos atributos e sua conexão com o problema;
2. **Ajustar a base de dados:** Os atributos são verificados 1 a 1, removendo aqueles que podem eventualmente atrapalhar na análise. Neste momento os atributos “timestamp” e “comments” foram excluídos manualmente, visto que é possível notar que são incapazes de ajudar na classificação do problema, ou até dificultar. Neste passo o atributo “gender” foi ajustado, i.e., foi filtrado e reagrupado usando um código em Python.
3. **Fazer a seleção de features:** É analisada as features mais relevantes a partir dos resultados de 4 métodos de busca bioinspirados: AntSearch (Colônia de formigas), CuckooSearch (Busca cuco), GeneticSearch (Algoritmo genético) e PSOSearch (Enxame de partículas). E a seleção foi feita para os atributos considerados relevantes para pelo menos metade dos métodos de busca citados, usando o mínimo de 50% de relevância.
4. **Testar os classificadores:** Foram usados os seguintes classificadores: Multilayer Perceptron, Random Forest, Random Tree, SVM, Naive Bayes, J48 e Redes Bayesianas, mudando seus hiperparâmetros para tentar encontrar as melhores configurações possíveis. Essa etapa se subdividiu em outras três:
 - a. Inicialmente foi aplicada a classificação usando cross-validation com 10 folds com a base já ajustada e, para cada algoritmo citado anteriormente, todos os hiperparâmetros foram ajustados a fim de encontrar os melhores resultados possíveis (Experimento 1);
 - b. Depois de encontrar o modelo para o problema, o aplicamos usando o dataset original (Experimento 2);
 - c. A partir disso e encontrando o “melhor método”, foi treinado 1 modelo com divisão de $\frac{2}{3}$ (dois terços) para treinamento e $\frac{1}{3}$ (um terço) para testes, para verificar como um modelo teoricamente irá desempenhar em testes reais do cotidiano (Experimento 3).
5. **Comparação das técnicas:** os resultados foram comparados e colocados sob análise para entender a influência dos hiperparâmetros no problema, assim como a influência dos métodos de busca na seleção de features.

III. RESULTADOS E DISCUSSÃO

A. Atributo *Gender*

Na base original, o atributo *Gender* tinha como instâncias mais de 46 classes, porque o método de coleta dos gêneros dos participantes foi um campo de texto, ao invés de opções selecionáveis, o que, além de ter o problema de erros de digitação, alguns usuários digitaram textos que sequer podem trazer alguma informação útil, como “Nah”, “A little about you”, “p”, entre outros. Dessa forma, foi necessário o reajuste desse atributo, inicialmente excluindo manualmente algumas instâncias que eram impossíveis de classificar, como as citadas anteriormente, e posteriormente usando um script em python, reclassificando as instâncias da seguinte maneira:

- Cis man
- Cis woman
- Trans man
- Trans woman
- Non binary
- Others

No.	Label	Count	Weight
1	Cis woman	247	247
2	Cis man	987	987
3	Trans woman	4	4
4	Non binary	5	5
5	Other	4	4

Fig. 1 Reorganização das classes do atributo *Gender*

Através da análise dos participantes da pesquisa com as novas classes neste dataset, foi possível determinar várias características relacionadas à população. A maioria dos participantes era do sexo masculino (aproximadamente 79,14%) com o restante identificando como mulheres cis (19,74%), mulheres trans (0,32%), outras e não binárias (0,40%) e homens trans (0%). No caso do gênero em pode ser que Mulher Trans, outros e não-binários não há muita influência no resultado obtido, já que o tamanho amostral não é estatisticamente relevante.

B. Seleção de features

Neste momento, é analisado as features mais relevantes a partir dos resultados de 4 métodos de busca bioinspirados:

- AntSearch (Colônia de formigas);
- CuckooSearch (Busca cuco);
- GeneticSearch (Algoritmo genético) e,
- PSOSearch (Enxame de partículas).

A seleção foi feita para os atributos considerados relevantes para pelo menos metade dos métodos de busca citados, usando o mínimo de 50% de relevância, ou seja, aquele atributo foi considerado importante por pelo menos metade dos *folds*.

Weka Explorer

Preprocess Classify Cluster Associate **Select attributes** Visualize

Attribute Evaluator
Choose **ClassifierSubsetEval** -B weka.classifiers.trees.RandomTree -T -H "Click to set hold out or test instances" -E DEFAULT -- -K 0 -M

Search Method
Choose **AntSearch** -N 20 -I 20 -T 0 -M 0.01 -A 0.9 -B 2.0 -C 0.7 -O 0 -X 0 -W 0 -Y 0 -Z 4.0 -V 0 -P 1 -R 20 -S 1 -L "C:\\Program Files\\Weka

Attribute Selection Mode
☐ Use full training set
☒ Cross-validation Folds
Seed

(Nom) treatment

Start Stop

Result list (right-click for options)

- 21:52:59 - AntSearch + ClassifierSubsetEval
- 21:53:22 - AntSearch + ClassifierSubsetEval
- 21:55:59 - AntSearch + ClassifierSubsetEval
- 21:57:14 - AntSearch + ClassifierSubsetEval

Attribute selection output

treatment

Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), s

number of folds (%)	attribute
10(100 %)	1 Age
10(100 %)	2 Gender
2(20 %)	3 Country
10(100 %)	4 state
2(20 %)	5 self_employed
0(0 %)	6 family_history
0(0 %)	7 work_interfere
10(100 %)	8 no_employees
1(10 %)	9 remote_work
0(0 %)	10 tech_company
0(0 %)	11 benefits
6(60 %)	12 care_options
9(90 %)	13 wellness_program
7(70 %)	14 seek_help
10(100 %)	15 anonymity
0(0 %)	16 leave
1(10 %)	17 mental_health_consequence
0(0 %)	18 phys_health_consequence
0(0 %)	19 coworkers
10(100 %)	20 supervisor
0(0 %)	21 mental_health_interview
8(80 %)	22 phys_health_interview
10(100 %)	23 mental_vs_physical
2(20 %)	24 obs_consequence

Status
OK

Log x 0

Fig. 2 Resultados de atributos considerados mais relevantes usando o método AntSearch

Weka Explorer

Preprocess Classify Cluster Associate **Select attributes** Visualize

Attribute Evaluator
Choose **ClassifierSubsetEval** -B weka.classifiers.trees.RandomTree -T -H "Click to set hold out or test instances" -E DEFAULT -- -K 0 -M

Search Method
Choose **CuckooSearch** -N 20 -I 20 -T 0 -M 0.01 -A 0.69657 -B 0.25 -O 0 -X 0 -Y 0 -Z 4.0 -V 0 -W 0 -P 1 -R 20 -S 1 -L "C:\\Program Files\\

Attribute Selection Mode
☐ Use full training set
☒ Cross-validation Folds
Seed

(Nom) treatment

Start Stop

Result list (right-click for options)

- 21:52:59 - AntSearch + ClassifierSubsetEval
- 21:53:22 - AntSearch + ClassifierSubsetEval
- 21:55:59 - AntSearch + ClassifierSubsetEval
- 21:57:14 - AntSearch + ClassifierSubsetEval
- 21:59:19 - CuckooSearch + ClassifierSubsetEval
- 21:59:36 - CuckooSearch + ClassifierSubsetEval**

Attribute selection output

treatment

Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), s

number of folds (%)	attribute
10 (100 %)	1 Age
8 (80 %)	2 Gender
1 (10 %)	3 Country
10 (100 %)	4 state
8 (80 %)	5 self_employed
0 (0 %)	6 family_history
2 (20 %)	7 work_interfere
10 (100 %)	8 no_employees
2 (20 %)	9 remote_work
0 (0 %)	10 tech_company
0 (0 %)	11 benefits
6 (60 %)	12 care_options
9 (90 %)	13 wellness_program
8 (80 %)	14 seek_help
6 (60 %)	15 anonymity
0 (0 %)	16 leave
1 (10 %)	17 mental_health_consequence
1 (10 %)	18 phys_health_consequence
1 (10 %)	19 coworkers
9 (90 %)	20 supervisor
1 (10 %)	21 mental_health_interview
8 (80 %)	22 phys_health_interview
9 (90 %)	23 mental_vs_physical
0 (0 %)	24 obs_consequence

Status
OK

Log x 0

Fig. 3 Resultados de atributos considerados mais relevantes usando o método CuckooSearch

Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Attribute Evaluator

Choose

ClassifierSubsetEval -B weka.classifiers.trees.RandomTree -T -H "Click to set hold out or test instances" -E DEFAULT -- -K 0 -M

Search Method

Choose

GeneticSearch -Z 20 -G 20 -C 0.6 -M 0.033 -R 20 -S 1

Attribute Selection Mode

Use full training set

Cross-validation

Folds

10

Seed

1

(Nom) treatment

Start

Stop

Result list (right-click for options)

21:52:59 - AntSearch + ClassifierSubsetEval

21:53:22 - AntSearch + ClassifierSubsetEval

21:55:59 - AntSearch + ClassifierSubsetEval

21:57:14 - AntSearch + ClassifierSubsetEval

21:59:19 - CuckooSearch + ClassifierSubsetEval

21:59:36 - CuckooSearch + ClassifierSubsetEval

22:01:17 - GeneticSearch + ClassifierSubsetEval

Attribute selection output

treatment

Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), s

number of folds (%)	attribute
10 (100 %)	1 Age
1 (10 %)	2 Gender
8 (80 %)	3 Country
10 (100 %)	4 state
0 (0 %)	5 self_employed
0 (0 %)	6 family_history
10 (100 %)	7 work_interfere
3 (30 %)	8 no_employees
0 (0 %)	9 remote_work
1 (10 %)	10 tech_company
0 (0 %)	11 benefits
1 (10 %)	12 care_options
8 (80 %)	13 wellness_program
0 (0 %)	14 seek_help
7 (70 %)	15 anonymity
2 (20 %)	16 leave
8 (80 %)	17 mental_health_consequence
7 (70 %)	18 phys_health_consequence
4 (40 %)	19 coworkers
1 (10 %)	20 supervisor
1 (10 %)	21 mental_health_interview
2 (20 %)	22 phys_health_interview
3 (30 %)	23 mental_vs_physical
7 (70 %)	24 obs_consequence

Status

OK

Log

x 0

Fig. 4 Resultados de atributos considerados mais relevantes usando o método GeneticSearch

Weka Explorer

Preprocess Classify Cluster Associate **Select attributes** Visualize

Attribute Evaluator
Choose **ClassifierSubsetEval** -B weka.classifiers.trees.RandomTree -T -H "Click to set hold out or test instances" -E DEFAULT -- -K 0 -M

Search Method
Choose **PSOSearch** -N 20 -I 20 -T 0 -M 0.01 -A 0.33 -B 0.33 -C 0.34 -R 20 -S 1 -L "C:\Program Files\Weka-3-9-6"

Attribute Selection Mode
☐ Use full training set
☒ Cross-validation Folds Seed

(Nom) treatment

Result list (right-click for options)

- 21:52:59 - AntSearch + ClassifierSubsetEval
- 21:53:22 - AntSearch + ClassifierSubsetEval
- 21:55:59 - AntSearch + ClassifierSubsetEval
- 21:57:14 - AntSearch + ClassifierSubsetEval
- 21:59:19 - CuckooSearch + ClassifierSubsetEval
- 21:59:36 - CuckooSearch + ClassifierSubsetEval
- 22:01:17 - GeneticSearch + ClassifierSubsetEval
- 22:02:35 - PSOSearch + ClassifierSubsetEval**

Attribute selection output

treatment
Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), s

number of folds (%)	attribute
10 (100 %)	1 Age
5 (50 %)	2 Gender
5 (50 %)	3 Country
10 (100 %)	4 state
1 (10 %)	5 self_employed
2 (20 %)	6 family_history
10 (100 %)	7 work_interfere
2 (20 %)	8 no_employees
2 (20 %)	9 remote_work
0 (0 %)	10 tech_company
1 (10 %)	11 benefits
3 (30 %)	12 care_options
6 (60 %)	13 wellness_program
6 (60 %)	14 seek_help
9 (90 %)	15 anonymity
0 (0 %)	16 leave
4 (40 %)	17 mental_health_consequence
1 (10 %)	18 phys_health_consequence
3 (30 %)	19 coworkers
1 (10 %)	20 supervisor
1 (10 %)	21 mental_health_interview
5 (50 %)	22 phys_health_interview
3 (30 %)	23 mental_vs_physical
1 (10 %)	24 obs_consequence

Status
OK x 0

Fig. 5 Resultados de atributos considerados mais relevantes usando o método PSOSearch

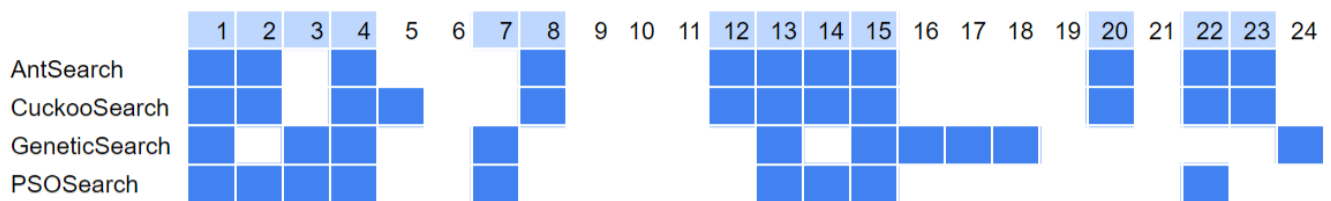


Fig.6 Mapeamento dos resultados dos métodos de busca por atributos mais relevantes. Os índices dos atributos destacados são os que foram escolhidos para análise.

Considerando que **treatment** é a classe que deve ser prevista (que indica caso a pessoa procurou ajuda para alguma condição mental) e, de acordo com a Figura 6, que mostra quais atributos foram selecionados por cada algoritmo, de acordo com o método descrito anteriormente, foram escolhidos, por fim, os seguintes atributos para análise:

- **Age: idade do participante.**

Um dos atributos esperados, já que a idade tem influência sobre algumas condições mentais (podem surgir mais em determinadas fases da vida).

- **Gender: gênero do participante.**

Outro atributo esperado, visto que alguns grupos minoritários podem se sentir um pouco mais excluídos e isto vir a dificultar suas jornadas de trabalho.

- **Country: país onde o participante mora.**

Não era um dos esperados, mas faz sentido, visto que países com variados IDHs podem ter taxas de condições mentais diferentes.

- **State: caso more nos EUA, seu estado ou território.**

Também não era esperado, mas pode ocorrer diferenças pelo motivo citado acima.

- **Work_Interference: se possui condição mental, ela interfere no trabalho?**

Não foi um dos considerados anteriormente, mas com certeza também pode influenciar, já que está conectado ao conforto provido pelo emprego.

- **No_employees: quantos empregados sua organização possui?**

Foi um atributo inusitado. É necessário maior investigação, mas algumas hipóteses estão ligadas a empresas maiores possuírem melhores condições de tratamentos ou, em empresas menores, o indivíduo se sentir mais confortável e poder ter maior proximidade com os superiores e colegas de trabalho.

- **Care_options: você conhece as opções de cuidados mentais fornecidos pela sua organização?**

Não foi um dos considerados, porém pode, claro, influenciar, visto que conhecer as opções de cuidados mentais pode ajudar a pessoa a cuidar melhor de sua saúde mental.

- **Wellness_program: sua organização já discutiu sobre saúde mental como parte de programa de bem-estar**

Um dos considerados, por um motivo parecido com o de cima: ter opções de cuidados mentais numa empresa estimula os funcionários a cuidar de sua saúde.

- **Seek_help: sua organização provê recursos para aprender mais sobre condições mentais e como procurar ajuda?**

Não foi um dos pensados anteriormente, mas também pode ser importante pelo motivo citado acima.

- **Anonymity: seu anonimato está protegido se você optar por aproveitar os recursos de tratamento de saúde mental ou abuso de substâncias?**

Não foi um dos considerados, mas é possível ver como pode ser interessante: pode se tornar mais fácil para algumas pessoas procurarem ajuda.

- **Supervisor: você estaria disposto a discutir um problema de saúde mental com seu supervisor direto?**

Também não tinha sido considerado, mas pode ser importante pelo mesmo motivo citado acima.

- **Phys_health_interview: você falaria sobre um problema de saúde física com um supervisor em uma entrevista?**

Também não tinha sido considerado, mas é um ponto interessante, visto que influencia na vida após conseguir o emprego.

- **Mental_vs_physical: você sente que seu chefe leva saúde mental tão sério quanto saúde física?**

Um dos que pensamos anteriormente, porque pode, por exemplo, significar que uma empresa não leva muito a sério o problema de saúde mental.

E excluídos os seguintes atributos:

- **Timestamp (excluído manualmente): quando a coleta foi realizada**

Não possui muita contribuição para o problema visto que todas as instâncias são únicas.

- **Comments (excluído manualmente): comentários adicionais**

Mesmo motivo acima.

- **Self_employed: se é autônomo**

Como esperado, acreditamos que ser autônomo possivelmente não muda significativamente o estado mental da pessoa.

- **Family_history: se tem algum caso de condição mental na família**

Era um dos esperados, já que acreditávamos que a genética poderia trazer consigo alguns genes que poderiam influenciar no indivíduo. Talvez neste caso não possua informações suficientes para que isto possa ser estimado ou a influência seja baixa.

- **Remote_work: você trabalha remotamente pelo menos 50% do tempo?**

Assim como esperado, é possível que trabalho remoto possa influenciar nas condições mentais, mas talvez essa influência seja um pouco menor que para outros atributos.

- **Tech_company: sua organização é principalmente de tecnologia?**

Outro atributo que consideramos que pode influenciar, mas sua influência pode não ser significativa neste problema.

- **Benefits: seu empregador possui benefícios para cuidar da saúde mental?**

Embora não tenha sido um dos considerados mais importantes, é interessante notar que não foi considerado importante, mesmo que “Care_options” e “Wellness_program” tenha sido.

- **Leave: o quão fácil é pra você tirar licença médica por uma condição de saúde mental?**

Acreditamos que poderia ter influência neste problema, como era o pensado anteriormente, mas talvez não seja tão significativo como outras características.

- **Mental_health_consequence: você acha que discutir um problema de saúde mental com seu superior traria consequências negativas?**

Acreditamos que possa ter influência neste problema, mas possivelmente sua significância deva ser menor que outras características.

- **Phys_health_consequence: você acha que discutir um problema de saúde física com seu superior traria consequências negativas?**

Acreditamos que possa ter influência neste problema, mas possivelmente sua significância deva ser menor que outras características.

- **Coworkers: você estaria disposto a discutir um problema de saúde mental com seus colegas de trabalho?**

É possível que se sentir confortável para discutir um problema de saúde mental com os colegas possa indicar um bem-estar no trabalho, mas que sua influência sobre a saúde mental não seja tão considerável quanto opções para cuidar da saúde mental, por exemplo.

- **Mental_health_interview: você falaria sobre um problema de saúde mental com um supervisor em uma entrevista?**

Acreditamos que talvez não tenha uma significância tão grande para este problema, visto que não indica com tanta precisão que uma pessoa possui um problema de saúde mental.

- **Obs_consequence: você já ouviu falar ou observou consequências negativas em colegas de trabalho com problemas de saúde mental em seu local de trabalho?**

Não foi considerada relevante inicialmente, mas com certeza influencia sim na saúde mental, porque significa que a organização talvez não esteja providenciando a ajuda necessária para o colega que esteja sofrendo com algum problema de saúde mental, e isso impacta todos os colaboradores.

C. Teste para o Experimento 1

Tabela 1 Resultados dos classificadores no Exp 1

Índice	Classificadores	Configurações dos hiperparâmetros	Acurácia		Kappa		Especificidade		Sensibilidade		Área da curva ROC	
			Média (%)	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
1	J48	Fator de confiança = 0.01, número mínimo de instâncias por folha = 2	83,00	2,91	0,66	0,06	0,71	0,05	0,95	0,03	0,86	0,03
2	J48	Fator de confiança = 0.1, número mínimo de instâncias por folha = 10	83,00	2,91	0,66	0,06	0,71	0,05	0,95	0,03	0,86	0,03
3	J48	Fator de confiança = 0.1, número mínimo de instâncias por folha = 2	82,96	2,90	0,66	0,06	0,71	0,05	0,95	0,03	0,86	0,03
4	J48	Fator de confiança = 0.25, número mínimo de instâncias por folha = 10	82,85	2,95	0,66	0,06	0,72	0,05	0,94	0,03	0,86	0,03
5	RandomForest	Porcentagem do tamanho do conjunto de treinamento = 100, tamanho da Floresta = 100, número de atributos escolhidos aleatoriamente = 0, profundidade da árvore = 3	81,95	2,90	0,64	0,06	0,74	0,05	0,90	0,04	0,87	0,03
6	RandomTree	Número de atributos escolhidos aleatoriamente = 13, mínimo total do peso das instâncias por folha = 50.0, profundidade máxima = 3	81,88	2,93	0,64	0,06	0,72	0,05	0,92	0,03	0,84	0,03
7	SMO	c = 1.0, parâmetro de tolerância = 0.001, epsilon = 1.0E-12, kernel = polinomial, expoente = 1.0 (linear); calibrador = Logistic, valor de Ridge na probabilidade de log = 1.0E-8, número máximo de iterações = até a convergência	81,94	2,74	0,64	0,06	0,76	0,06	0,94	0,03	0,82	0,03
8	SMO	c = 0.1, parâmetro de tolerância = 0.001, epsilon = 1.0E-12, kernel = polinomial, expoente = 1.0 (linear); calibrador = Logistic, valor de Ridge na probabilidade de log = 1.0E-8, número máximo de iterações = até a convergência	81,54	2,65	0,62	0,05	0,71	0,06	0,92	0,03	0,87	0,08
9	BayesNet	Algoritmo de busca = K2 (número máximo de pais por nó = 1, tipo de pontuação = BAYES), estimador = SimpleEstimator, alfa = 0.5	79,95	3,22	0,60	0,06	0,76	0,05	0,84	0,05	0,87	0,03
10	NaiveBayes	usando estimador kernel	79,73	3,16	0,59	0,06	0,76	0,05	0,84	0,04	0,87	0,03
11	NaiveBayes	-	79,51	3,25	0,59	0,07	0,75	0,05	0,84	0,05	0,87	0,03
12	MLP	Taxa de aprendizado = 0.3, momentum = 0.2, épocas = 500, porcentagem do conjunto de validação = 0%, paciência = 20, neurônios (1 camada) = 100	79,01	4,13	0,58	0,08	0,77	0,07	0,81	0,11	0,86	0,03
13	RandomTree	Número de atributos escolhidos aleatoriamente = 13, mínimo total do peso das instâncias por folha = 1.0, máximo de profundidade da árvore = 3	78,66	3,16	0,57	0,06	0,71	0,05	0,87	0,04	0,81	0,04
14	MLP	Taxa de aprendizado = 0.3, momentum = 0.2, épocas = 500, porcentagem do conjunto de validação = 0%, paciência = 20, neurônios (1 camada) = 50	77,84	3,79	0,56	0,08	0,77	0,07	0,79	0,10	0,86	0,03
15	RandomTree	Número de atributos escolhidos aleatoriamente = 13, mínimo total do peso das instâncias por folha = 1.0, profundidade máxima = ilimitado	75,15	3,44	0,50	0,07	0,75	0,05	0,76	0,05	0,75	0,04

No experimento acima já é possível notar algumas características interessantes:

- J48 conseguiu melhores resultados que todos os outros modelos, incluindo Random Forest e MLP, que são modelos considerados “fortes”.
 - No caso do Random Forest, mudar a profundidade máxima não mostrou melhoras nas profundidades de 3, 5 e 10 mostraram resultados de 81,54%, 81,26% e 81,16%, respectivamente. E aumentar o tamanho da floresta também não ajudou, pelo contrário, piorou os resultados. Uma possível justificativa é de que pode ter dificultado a generalização do modelo. Neste caso, aumentar de 100 para 500 diminuiu de 81,95% para 80,72%, e para 1.000, diminuiu para 79,85%;
 - No caso do MLP infelizmente não foi possível testar com mais de uma camada por dificuldades no software utilizado (WEKA), o que pode ter influenciado significativamente sua performance, visto que impossibilitaria de gerar funções mais complexas.
- SVM apresentou resultados melhores que MLP, o que pode ser explicado pelo motivo citado anteriormente com relação à quantidade de camadas utilizadas;
- Naive Bayes e BayesNet apresentaram, também resultados melhores que MLP com apenas 1 camada, possivelmente pelo motivo anterior;
- O melhor resultado para Random Forest e Random Tree apresentaram valores bem próximos;

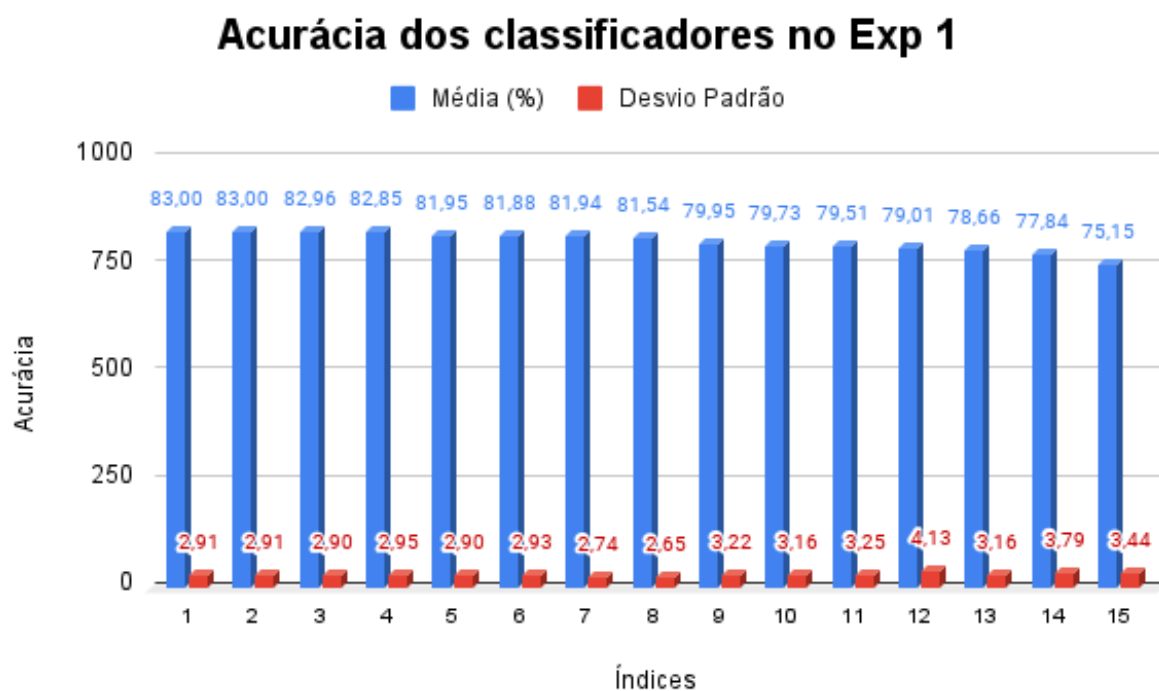


Fig. 7 Resultados das acurácias dos classificadores no Experimento 1

Como visto anteriormente e partir do gráfico mostrado na Figura 7, é possível visualizar que as quatro melhores classificações foram obtidas usando J48, nas configurações descritas na Tabela 1, com a melhor configuração de MLP ocupando a pior posição. No caso do MLP sua acurácia baixa pode ser explicada pelo motivo citado anteriormente de dificuldade em utilizar mais de 1 camada.

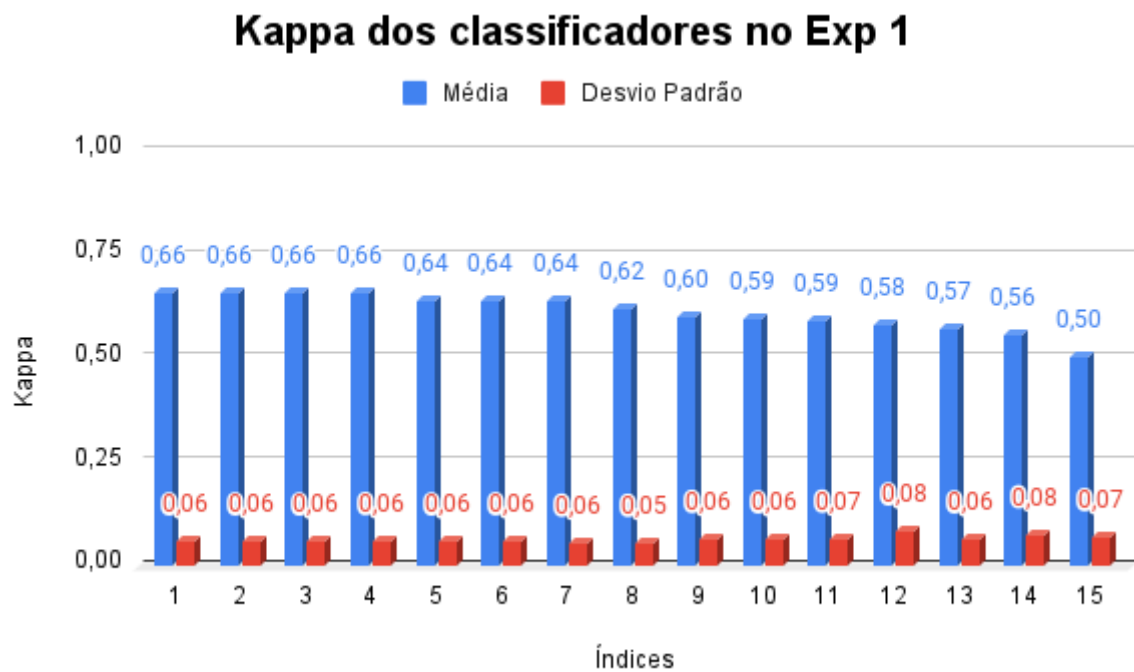


Fig. 8 Resultados do Kappa dos classificadores no Experimento 1

Da mesma forma, os quatro maiores Kappas foram obtidos usando J48, nas configurações descritas na Tabela 1, seguindo uma tendência bem próxima ao gráfico anterior.

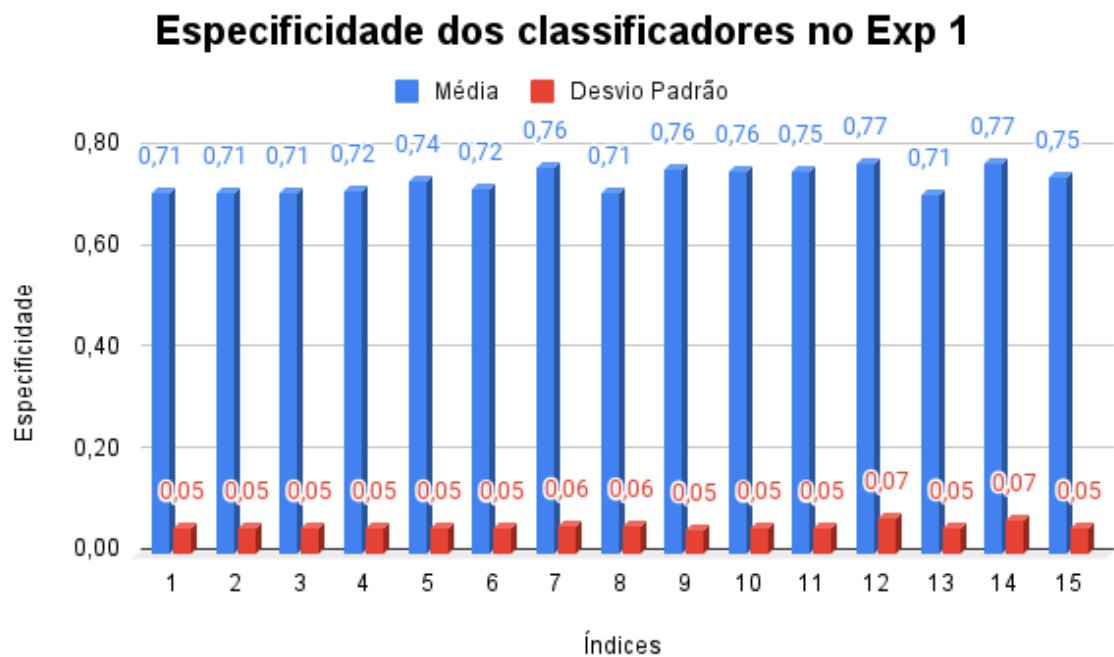


Fig. 9 Resultados das especificidades dos classificadores no Experimento 1

Com relação às especificidades, ou seja, às taxa de verdadeiros negativos, apresentaram os melhores (menores) resultados os testes enumerados 1, 2, 3, 4, 6, 8 e 13, ou seja, usando J48, Random Tree e um dos testes usando SMO.

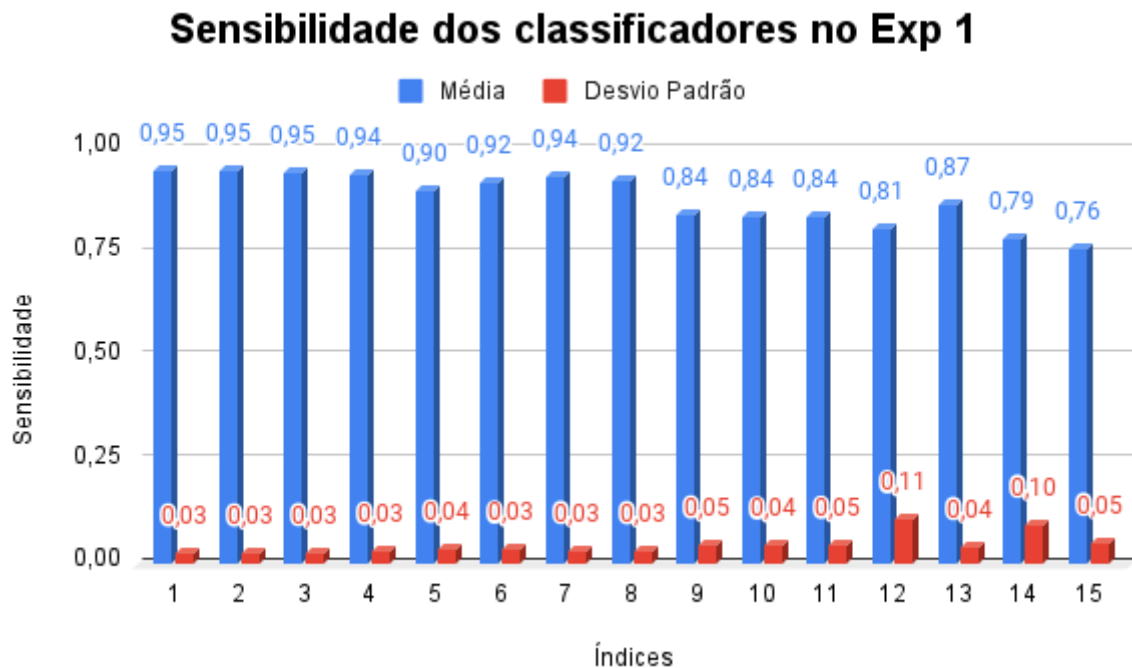


Fig. 10 Resultados das sensibilidades dos classificadores no Experimento 1

A respeito da sensibilidade, os melhores resultados foram alcançados usando J48 e uma das configurações usando SMO. Esta é a métrica considerada mais importante para o problema, visto que, embora todas sejam importantes, é importante que o método seja capaz de acertar com precisão quando uma pessoa pode vir ou está a vir a sofrer algum problema. Embora errar e considerar que uma pessoa está doente possa ser um problema, ainda, para este problema, um problema menor, visto que poderá vir a ter um cuidado maior com aquele indivíduo, porém deixar passar despercebido uma pessoa que precisa de cuidados seria péssimo.

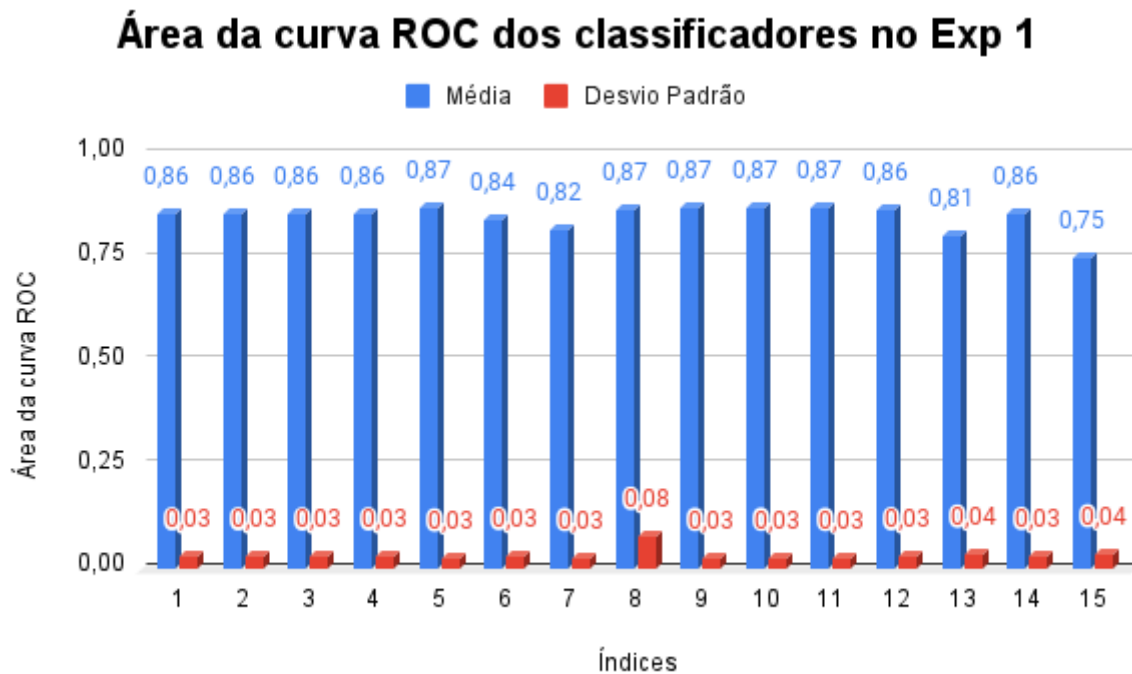


Fig. 11 Resultados das áreas da curva ROC dos classificadores no Experimento 1

Por fim, quanto à área da curva ROC, dos 15 experimentos apresentados, 14 ficaram acima de 0.8, mostrando resultados satisfatórios para o problema.

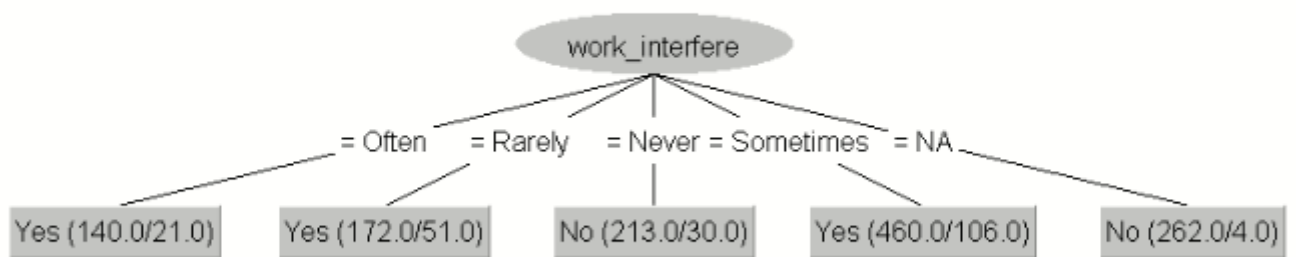


Fig. 12 Visualização da árvore de decisão

Por fim, foi visualizado a árvore de decisão do melhor dos algoritmos testados, o que surpreendeu: apenas utilizando o atributo “work_interfere” ele foi capaz de estimar o resultado com 83% de acurácia. Isto mostra que apenas verificar se, caso a pessoa possua alguma condição mental, ela ser capaz de interferir no trabalho é um fator extremamente importante para entender como melhorar a vida do profissional, mostrando que melhorar este aspecto pode não só melhorar sua qualidade de vida, mas seu desempenho no trabalho.

D. Teste para o Experimento 2

Tabela 2 Resultados dos classificadores no Exp 2

Índice	Classificadores	Configurações dos hiperparâmetros	Acurácia		Kappa		Especificidade		Sensibilidade		Área da curva ROC	
			Média (%)	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
1	J48	Fator de confiança = 0.01, número mínimo de instâncias por folha = 2	83,00	2,91	0,66	0,06	0,71	0,05	0,95	0,03	0,86	0,03
2	J48	Fator de confiança = 0.25, número mínimo de instâncias por folha = 5	82,99	2,90	0,66	0,06	0,71	0,05	0,95	0,03	0,86	0,03
3	J48	Fator de confiança = 0.25, número mínimo de instâncias por folha = 10	82,96	2,92	0,66	0,06	0,71	0,05	0,94	0,03	0,86	0,03
4	J48	Fator de confiança = 0.1, número mínimo de instâncias por folha = 2	82,80	2,97	0,66	0,06	0,71	0,05	0,94	0,03	0,86	0,03
5	J48	Fator de confiança = 0.25, número mínimo de instâncias por folha = 2	82,52	2,93	0,65	0,06	0,72	0,05	0,93	0,04	0,86	0,03

A partir dos resultados do Experimento 2 mostrados na Tabela 2, é possível inferir que:

- Os valores de acurácia dos três primeiros testes foram muito próximos;
- Otimizando os hiperparâmetros para a base com todos os atributos, temos um resultado bem próximo do Experimento 1, com as features selecionadas, o que sugere que pouca ou nenhuma informação foi perdida. Dessa forma, poderíamos utilizar o outro modelo, tornando mais simples o algoritmo, tanto em eficiência, quanto em tempo de execução e, em caso de árvores, visualização.

Acurácia dos classificadores no Exp 2

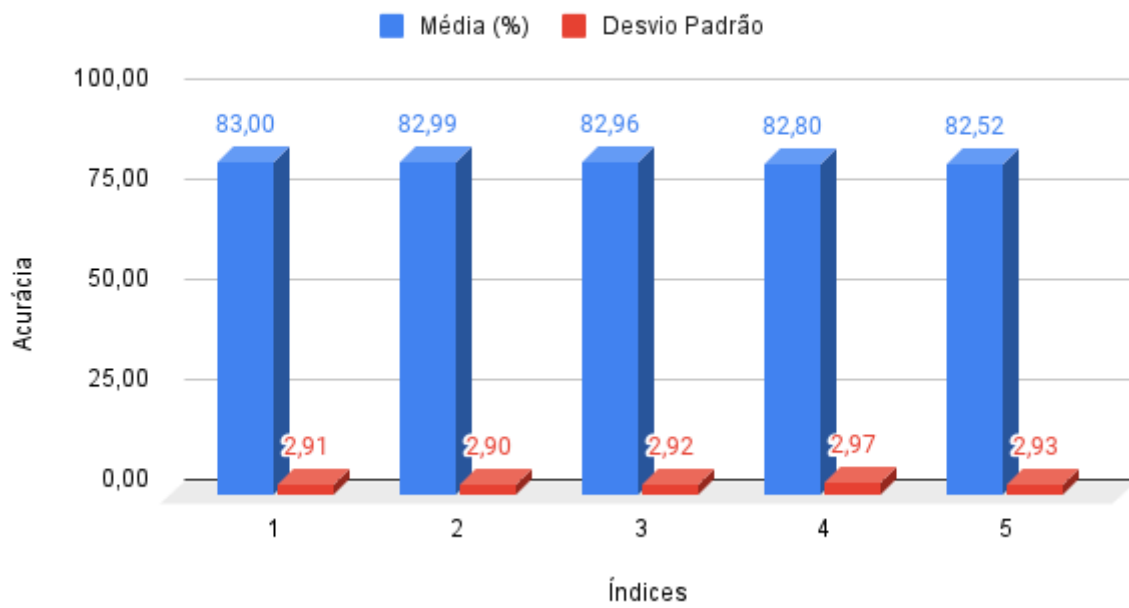


Fig. 13 Resultados das acurácias dos classificadores no Experimento 2

Em relação às acurácias no Experimento 2, foram atingidos melhores resultados nos três primeiros testes.

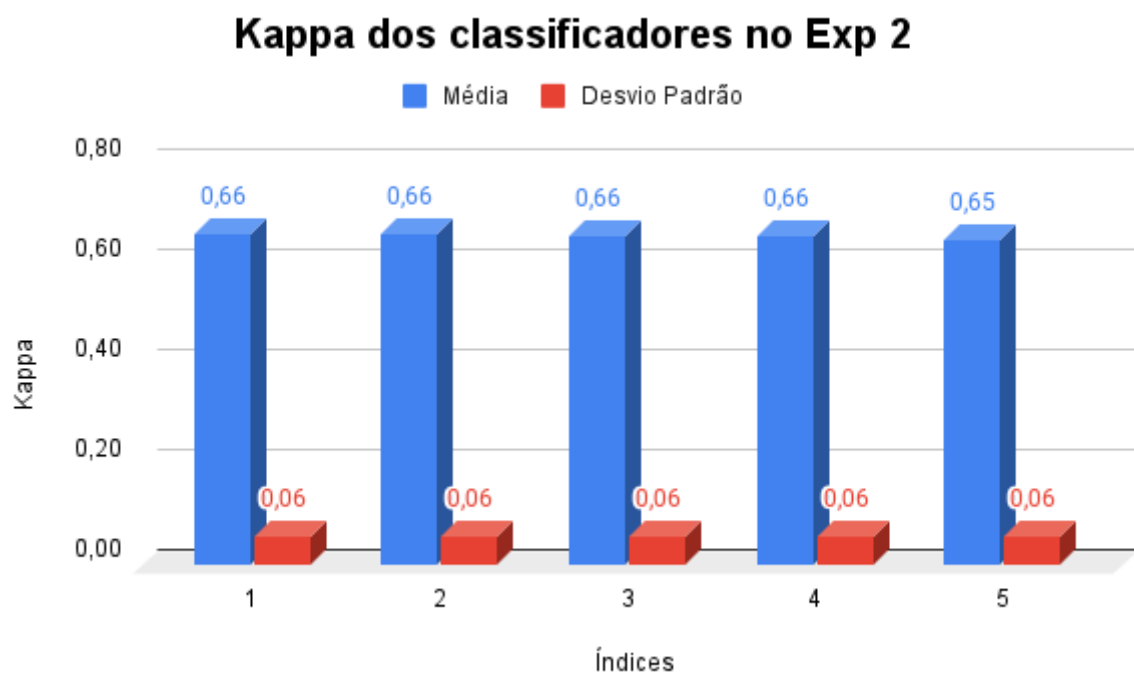


Fig. 14 Resultados do Kappa dos classificadores no Experimento 2
Sobre o Kappa, os valores foram praticamente iguais para todos os cinco testes.

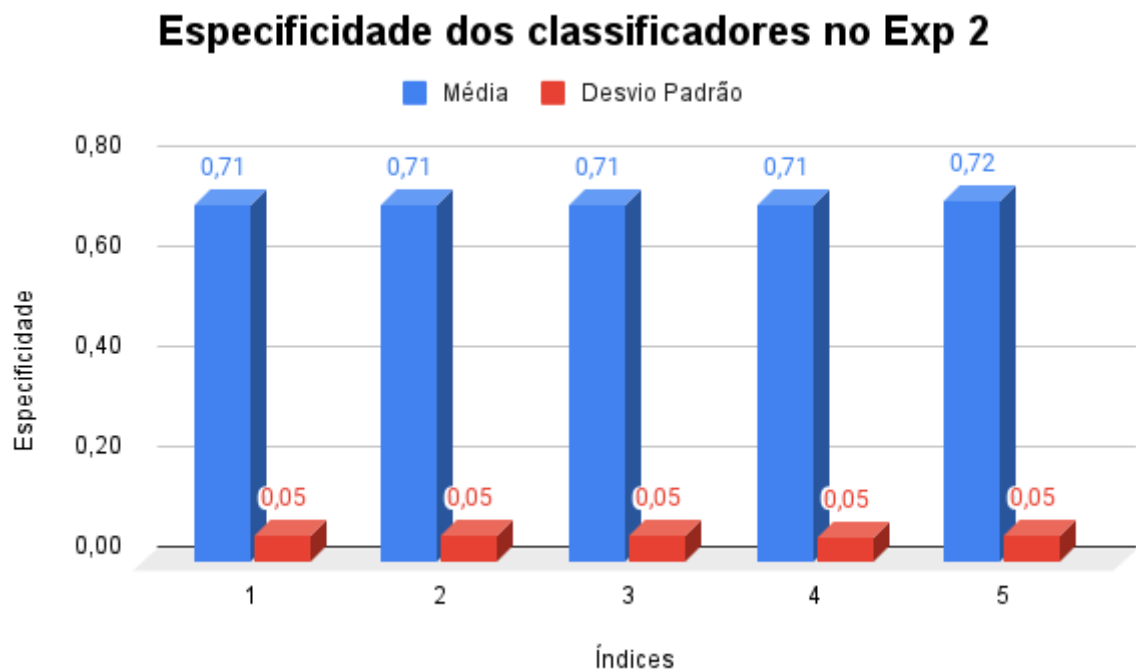


Fig. 15 Resultados das especificidades dos classificadores no Experimento 2
Com relação à especificidade, os resultados foram igualmente próximos para todos os testes.

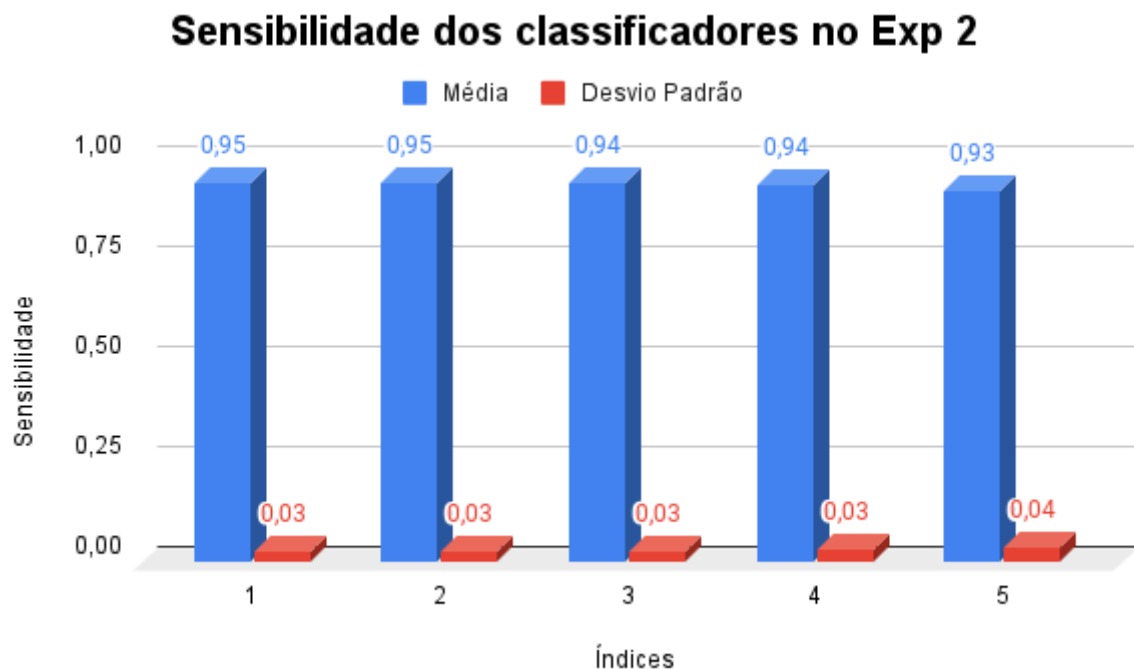


Fig. 16 Resultados das sensibilidades dos classificadores no Experimento 2
Essa proximidade e semelhança também pode ser observada na sensibilidade das classificações.

Área da curva ROC dos classificadores no Exp 2

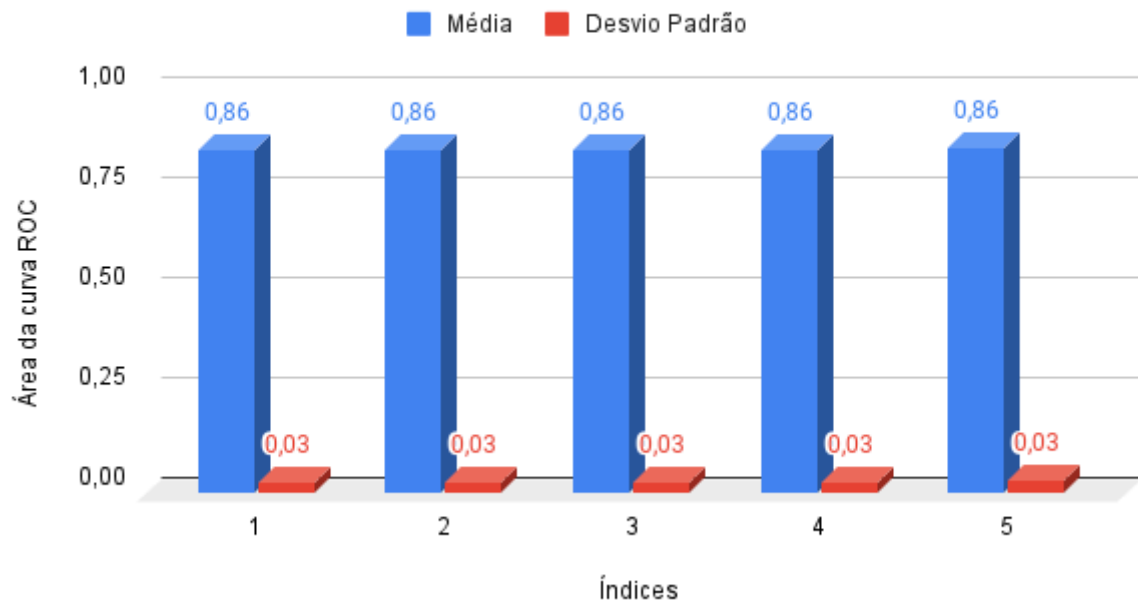


Fig. 17 Resultados da área da curva ROC dos classificadores no Experimento 2
Como previsto, a área da curva ROC foi exatamente a mesma para todos os cinco testes.

E. Teste para o Experimento 3

Tabela 3 Resultados dos classificadores no Exp 3

Classificadores	Configurações dos hiperparâmetros	Acurácia	Kappa	Especificidade	Sensibilidade	Área da curva ROC
J48	Fator de confiança = 0.01, número mínimo de instâncias por folha = 2	0,8137	0,6269	0,6730	0,9531	0,8424

Sobre o Experimento 3, pode-se concluir que:

- Os valores obtidos para todas as métricas foram bem próximos dos valores dos outros experimentos, em que foi aplicado cross-validation de 10 *folds*. Dessa forma, é possível inferir que o modelo obtido pode ser considerado fiel quando aplicado a problemas reais.

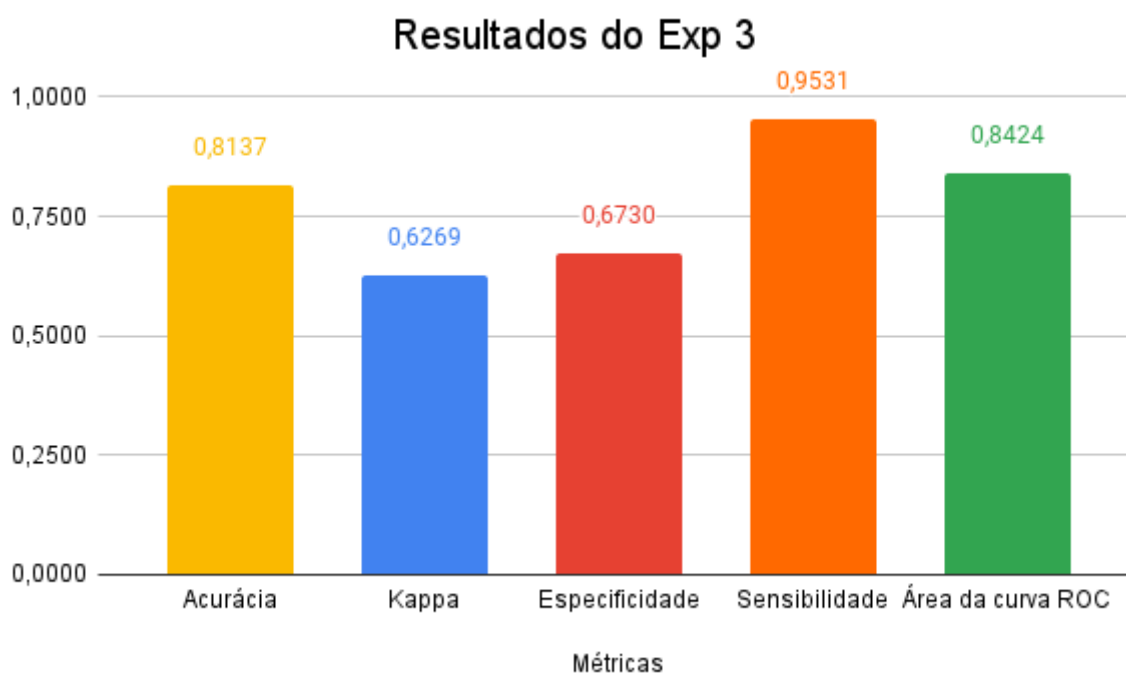


Fig. 18 Resultados das métricas no Exp 3

IV. CONCLUSÃO

Para concluir, o principal entendimento que extraímos deste projeto é que parece haver uma associação direta entre tratamento de saúde mental e idade dos funcionários, gênero, país, estado, interferência no trabalho, número de funcionários, opções de cuidados, programa de bem-estar, procurar ajuda, anonimidade, supervisor, opções de cuidados no local de trabalho e tratamento. Isso forneceu evidências para nossa questão de pesquisa por meio da exploração e modelagem de dados. Por meio do software Weka, selecionamos um conjunto de dados de alta qualidade e nossas técnicas de pré-processamento ajudaram com sucesso na construção de um modelo que pode prever recursos no conjunto de dados.

Dessa forma e a partir dos experimentos realizados, pode-se concluir que as hipóteses foram sim respondidas. Com relação ao atributo *Gender*, pode-se constatar com base nos dados obtidos que os homens cis são as mais afetadas pelo problema de saúde mental na tecnologia. Além disso, observa-se que a pesquisa teve um percentual maior de homens cis do que qualquer outra classificação de gênero. O fato de haver um maior percentual de homens no estudo afeta os resultados obtidos. Isso vai de encontro à hipótese original em relação a esse conjunto de dados, no entanto, isso não quer dizer que as mulheres cis ou pessoas transgêneras, não-binárias e de outros gêneros não sejam afetadas. Existe a possibilidade de que, durante o período em que esses dados foram coletados, a maioria da indústria de tecnologia era de fato do sexo masculino, o que, como consequência, resultou em um número maior de participantes do sexo masculino nesta pesquisa.

Sobre o desequilíbrio dos dados com relação aos estados e países, pode-se notar que durante o tempo em que este estudo foi feito a maioria dos participantes eram pessoas nos Estados Unidos da América, isso pode ter sido devido a vários fatores limitantes durante esse período. Para este conjunto de dados, os estados dos Estados Unidos são extremamente relevantes, pois nos ajudam a entender melhor onde nos Estados Unidos tem o maior número de indivíduos afetados, ao mesmo tempo que impede e limita a análise em outras localidades do mundo.

Acerca dos atributos que foram considerados relevantes para os participantes, tal como, Age, Gender, e mental_vs_physical foram classificados como relevantes na seleção de atributos. Os atributos que se esperava que fossem vistos como relevantes, como Family History, Tech Company, Wellness Program, Leave, Mental Health Consequence, não foram classificados como relevantes, no entanto, isso pode ser devido a vários motivos e não significa que esses atributos não tenham um efeito na saúde mental. Neste conjunto de dados específico esses atributos específicos não têm dados suficientes para resultar em mudanças significativas nos resultados.

No final deste relatório, discutimos os dados e classificadores, onde podemos ver como os classificadores e os dados estão relacionados entre si e dando o melhor possível, resultados precisos e precisos. Também pode ser visto como a frequência da doença de saúde mental e a atitude em relação à saúde mental variam de acordo com a localização geográfica, como país e estado e quais foram os preditores mais fortes dos distúrbios da saúde mental. No entanto, embora os resultados gerados não tenham sido exatamente os mesmos mencionados na hipótese, podemos ver a partir dos experimentos que foram feitos para no processo de tomada de decisão que os resultados obtidos usando Weka podem ajudar no domínio da saúde mental.

REFERÊNCIAS

- [1] <https://itforum.com.br/noticias/burnout-dispara-com-covid-19-mais-de-70-dos-profissionais-em-ti-estao-esgotados/>
- [2] <https://www.scielo.br/j/pcp/a/pVQSQX5SPLbyzXCdM8DZT7v/?lang=pt>
- [3] <https://extra.globo.com/noticias/saude-e-ciencia/burnout-1-cada-5-profissionais-de-grandes-corporacoes-sofre-de-esgotamento-no-brasil-mostra-pesquisa-inedita-25589274.html>
- [4] <https://www.bbc.com/portuguese/vert-cap-62685199#:~:text=Existem%20duas%20barreiras%20diferentes%2C%20mas,de%20falar%20sobre%20o%20assunto>
- [5] <https://www.sciencedirect.com/science/article/abs/pii/S027795361830121>