

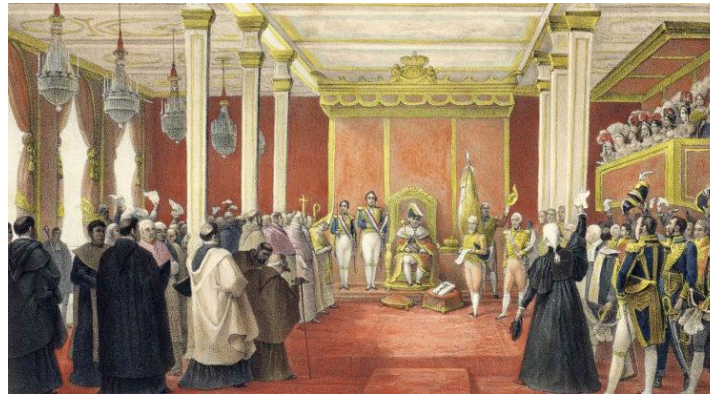
ANÁLISE DE PREVISIBILIDADE DO RESULTADO DAS ELEIÇÕES DE 2018 PARA OS CARGOS DE DEPUTADOS ESTADUAIS E FEDERAIS

Trabalho de Conclusão de Curso apresentado ao Curso de Especialização em
Ciência de Dados e Big Data – Henrique Moreira de Oliveira

HISTÓRIA DAS ELEIÇÕES NO BRASIL



1532



1808



1822



1889

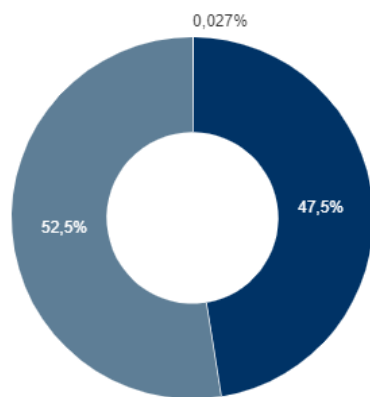


1964

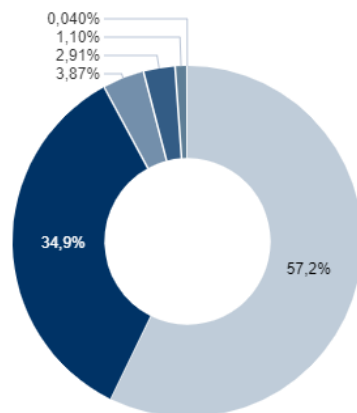
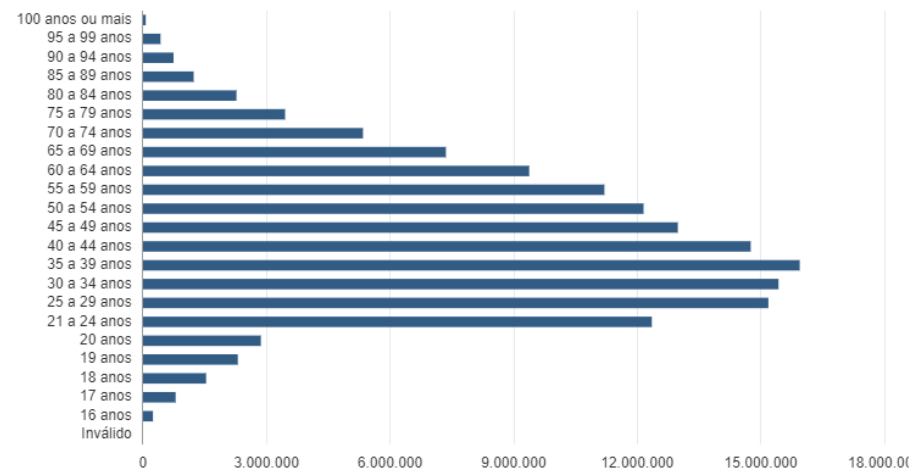


1988

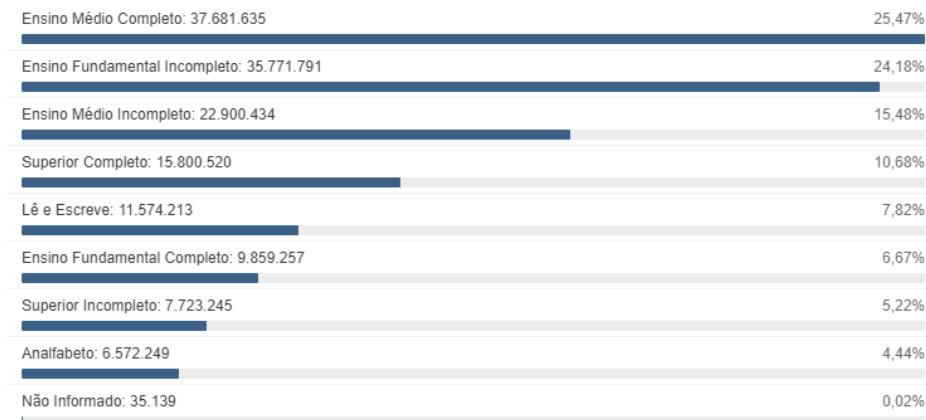
PERFIL DO ELEITORADO BRASILEIRO



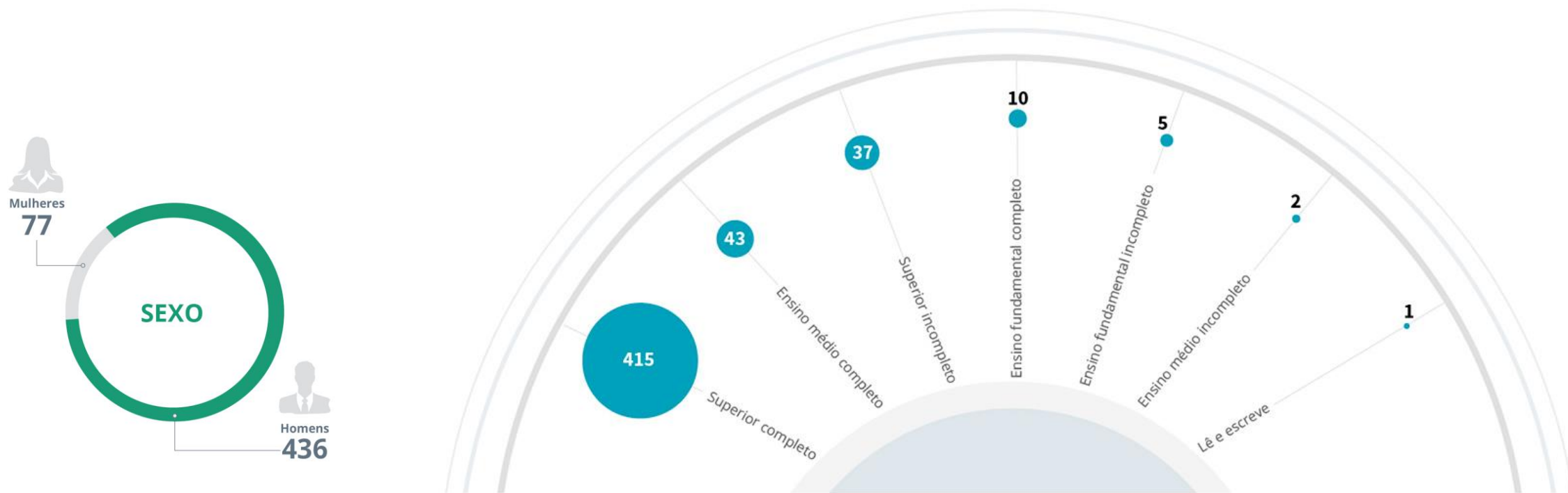
■ Não Informado
■ Masculino
■ Feminino



■ Solteiro
■ Casado
■ Divorciado
■ Viúvo
■ Separado Judicialmente
■ Não Informado



PERFIL DOS DEPUTADOS FEDERAIS



PROBLEMA PROPOSTO

Seria possível prever, com base no gênero, estado civil, idade, raça, escolaridade e poder aquisitivo dos candidatos, o resultado das eleições para os cargos de deputado federal e estadual na eleição de 2018, que é a última realizada até o momento?



Dados dos candidatos

• 29.145 entradas

Bens dos candidatos

• 93.271 entradas

Despesas contratadas

• 1.711.062 entradas

Despesas pagas

• 1.647.553 entradas

SOLUÇÕES ANALISADAS

Classificação é o processo de reconhecer, compreender e agrupar ideias e objetos em categorias predefinidas ou "subpopulações". Usando conjuntos de dados de treinamento pré-categorizados, os programas de aprendizado de máquina usam uma variedade de algoritmos para classificar conjuntos de dados futuros em categorias.



Árvore de
Decisão

Regressão
Logística

Naïve Bayes

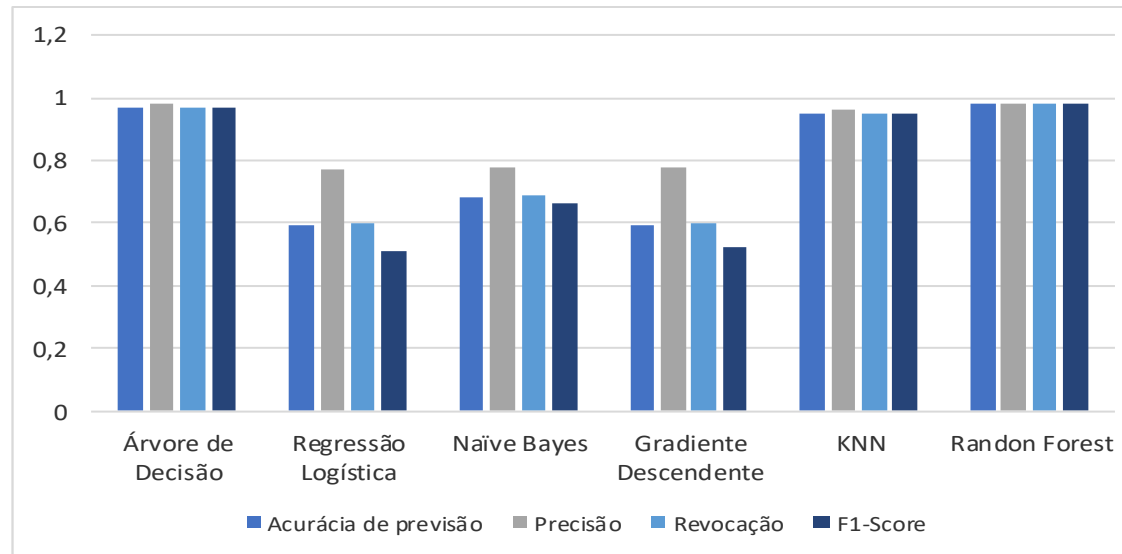
Gradiente
Descendente

KNN (K -
Nearest
Neighbors)

Randon
Forest

RESULTADOS

Algoritmo	Acurácia de previsão	Precisão	Revocação	F1-Score
Árvore de Decisão	0,97	0,98	0,97	0,97
Regressão Logística	0,59	0,77	0,60	0,51
Náïve Bayes	0,68	0,78	0,69	0,66
Gradiente Descendente	0,59	0,78	0,60	0,52
KNN	0,95	0,96	0,95	0,95
Randon Forest	0,98	0,98	0,98	0,98



CONCLUSÃO

O algoritmo Random Forest apresentou os melhores resultados, tendo atingido o valor de 0,98 para todas as medidas de avaliação (acurácia, precisão, revocação e F1-score)

Os resultados obtidos demonstram que o resultado da eleição é muito previsível, sendo possível afirmar com 98% de certeza se um candidato será ou não eleito.

Em resumo, se não existirem ações concretas para mudar o perfil dos candidatos e para dar melhores condições de vida à população, o perfil dos grupos que compõem as casas legislativas em todo o país continuará sendo o mesmo de 500 anos atrás, ou seja, homem branco pertencente à elite da sociedade.

The background is a complex network of thin grey lines connecting various sized nodes. The nodes are colored in dark blue, light blue, and grey. Some nodes are simple dots, while others are larger circles with concentric outlines. A large black rectangle is positioned in the lower right, containing the text 'OBRIGADO!' in white. A thin blue horizontal line is located just below the text.

OBRIGADO!