

VITOR ALEXANDRE KESSLER DE ALMEIDA

**PREDITOR DE ATENDIMENTO DAS RECOMENDAÇÕES
FORMULADAS PELA CONTROLADORIA-GERAL DA
UNIÃO**

Brasília

2020

VITOR ALEXANDRE KESSLER DE ALMEIDA

**PREDITOR DE ATENDIMENTO DAS RECOMENDAÇÕES
FORMULADAS PELA CONTROLADORIA-GERAL DA
UNIÃO**

Trabalho de conclusão do curso de pós-graduação
lato sensu em Análise de Dados para o Controle
realizado pela Escola Superior do Tribunal de
Contas da União como requisito para a obtenção do
título de especialista.

Orientador: Prof. Msc..Rodrigo Peres

Brasília

2020

REFERÊNCIA BIBLIOGRÁFICA

ALMEIDA, Vitor Alexandre Kessler de. **Preditor de atendimento das recomendações formuladas pela Controladoria-Geral da União**. 2020. Trabalho de Conclusão de Curso (Especialização em em Análise de Dados para o Controle) – Escola Superior do Tribunal de Contas da União, Instituto Serzedello Corrêa, Brasília, 2020.

CESSÃO DE DIREITOS

NOME DO AUTOR: Vitor Alexandre Kessler de Almeida

TÍTULO: Preditor de atendimento das recomendações formuladas pela Controladoria-Geral da União

GRAU/ANO: Especialista/2020

É concedido ao Instituto Serzedello Corrêa (ISC) permissão para reproduzir cópias deste Trabalho de Conclusão de Curso e emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. Do mesmo modo, o ISC tem permissão para divulgar este documento em biblioteca virtual, em formato que permita o acesso via redes de comunicação e a reprodução de cópias, desde que protegida a integridade do conteúdo dessas cópias e proibido o acesso a partes isoladas desse conteúdo. O autor reserva outros direitos de publicação e nenhuma parte deste documento pode ser reproduzida sem a autorização por escrito do autor.

Vitor Alexandre Kessler de Almeida
vitorkessler@gmail.com

Ficha catalográfica

Almeida, Vitor Alexandre Kessler. Preditor de atendimento das recomendações formuladas pela Controladoria-Geral da União / Vitor Alexandre Kessler de Almeida Almeida ; orientador, Rodrigo Peres, 2020. 52 p.

Trabalho de Conclusão de Curso (Especialização) Escola Superior do Tribunal de Contas da União, Instituto Serzedello Corrêa, Brasília, 2020. Inclui referências.

1. Classificação de Textos. 2. Controle Interno. 3. Recomendações. 4. Redes Neurais. 5. Aprendizado de Máquina.

VITOR ALEXANDRE KESSLER DE ALMEIDA

**PREDITOR DE ATENDIMENTO DAS RECOMENDAÇÕES
FORMULADAS PELA CONTROLADORIA-GERAL DA
UNIÃO**

Trabalho de conclusão do curso de pós-graduação lato sensu em Análise de Dados para o Controle realizado pela Escola Superior do Tribunal de Contas da União como requisito para a obtenção do título de especialista.

Local, xx de xxxx de xxxx.

Banca Examinadora:

Prof. Rodrigo Peres, MSc.

Orientador

Controladoria-Geral da União

Prof. Gustavo Cordeiro Galvao Van Erven, MSc.

Controladoria-Geral da União

AGRADECIMENTOS

Agradeço à Controladoria-Geral da União pela iniciativa em formar servidores da casa em Análise de Dados e por ter cedido a base de dados de recomendações necessária ao presente trabalho. Em especial, aos meus chefes nesse período de ausências nas segundas-feiras de curso e na licença capacitação. Foi por uma boa causa!

Agradeço ao ISC/TCU pelo pioneirismo em iniciar um curso de pós-graduação em análise de dados voltado para o controle; aos professores pelos ensinamentos repassados; e aos colegas de sala pelo companheirismo ao longo do curso e pelos momentos de diversão, em especial ao Pacote, nosso representante e colega de baia em todas as segundas-feiras do curso!

Agradeço aos colegas da CGU que fizeram o curso comigo pelos diversos trabalhos em equipe: o Claudio, o Rangel, a Renata e a Lucilene. São amigos que irei levar pro resto da vida!

Agradeço aos almoços de segunda-feira com meus colegas de curso, regados a muito café!

Agradeço a minha esposa pelo suporte familiar, pelo amor incondicional e pela dedicação, fundamentais nessa caminhada.

Agradeço a Deus, pelo seu amor incomensurável e por tudo o que ele é na minha vida.

RESUMO

O presente trabalho teve como objetivo desenvolver um classificador de texto das recomendações emitidas pela CGU, com vistas a identificar a viabilidade de atendimento de uma recomendação de auditoria pelo gestor público federal. Para tanto, foi seguida a metodologia CRISP-DM, sendo construídos modelos utilizando as seguintes técnicas de Aprendizado de Máquina: regressão logística, Naive Bayes, Support Vector Machine e redes neurais. O modelo que obteve melhor acurácia, acertando o atendimento ou não de 69% das recomendações utilizadas na avaliação foi aquele baseado em redes neurais. Os trabalhos futuros envolvem a implantação do modelo e a melhoria do pré-processamento de dados e das técnicas utilizadas.

Palavras-chave: Classificação de Textos. Controle Interno. Recomendações de Auditoria. Redes Neurais. Aprendizado de Máquina.

ABSTRACT

This work aimed to develop a text classifier of the recommendations issued by CGU, in order to identify the feasibility of conclude a recommendation by the federal public manager. For that, the CRISP-DM methodology was followed, with models being built using the following Machine Learning techniques: logistic regression, Naive Bayes, Support Vector Machine and neural networks. The model that obtained the best accuracy, hitting positively 69% of the recommendations, was based on neural networks Future work involves the implementation of the model and the improvement of data pre-processing and the techniques used.

Keywords: Text Classification. Internal control. Audit Recommendations. Neural networks. Machine Learning.

LISTA DE ILUSTRAÇÕES

Figura 1: Etapas do Modelo CRISP-DM.	18
Figura 2: Comparação entre a regressão linear e a regressão logística.....	34
Figura 3: Funcionamento de uma SVM, mostrando a separação das classes por meio de hiperplanos.....	35
Figura 4: Representação esquemática de um perceptron de uma rede neural.	36
Figura 5: Representação esquemática de uma rede neural MLP.	36
Figura 6: Representação esquemática do processo de avaliação dos modelos utilizados no presente trabalho.	38
Figura 7: Representação esquemática de uma matriz de confusão.....	39
Figura 8: Precisão, <i>recall</i> e <i>f1-score</i> obtidos na etapa de testes do melhor modelo de Regressão Logística.....	45
Figura 9 - Precisão, <i>recall</i> e <i>f1-score</i> obtidos na etapa de testes do melhor modelo de Multinomial Naive Bayes.	46
Figura 10 - Precisão, <i>recall</i> e <i>f1-score</i> obtidos na etapa de testes do melhor modelo de <i>Support Vector Machine</i>	47
Figura 11 - Precisão, <i>recall</i> e <i>f1-score</i> obtidos na etapa de testes do melhor modelo de Rede Neural.....	48

LISTA DE TABELAS

Tabela 1: Compilação de trabalhos de text classification em textos em português utilizando técnicas de ML. Resultados ordenados pela precisão média do melhor modelo encontrado. .41	
Tabela 2 - Resultados obtidos com a variação de parâmetros dos modelos treinados, obtidos por meio de validação cruzada 10-k fold.43	
Tabela 3: Comparação da precisão e <i>recall</i> médios (<i>micro avg</i>) dos modelos desenvolvidos no presente trabalho.48	

LISTA DE ABREVIATURAS E SIGLAS

AI: *Artificial Intelligence*

BOW: *Bag of Words*

CGU: Controladoria-Geral da União

COSO: *Committee of Sponsoring Organizations of the Treadway Commission*

CRISP-DM: *Cross-Industry Standard Process for Data Mining*

csv: *comma-separated values*

ML: *Machine Learning*

NB: *Naive Bayes*

PLN: Processamento de Linguagem Natural

PPP: Plano de Providências Permanente

RL: Regressão Logística

SCI: Sistema de Controle Interno

SFC: Secretaria Federal de Controle Interno

SGD: *Stochastic Gradient Descendent*

SVM: *Support Vector Machines*

TF-IDF: *Term Frequency - Inverse Document Frequency*

TCU: Tribunal de Contas da União

SUMÁRIO

1	INTRODUÇÃO.....	15
2	OBJETIVOS.....	17
2.1	Objetivo Geral.....	17
2.2	Objetivo Específico.....	17
3	DESENVOLVIMENTO	18
3.1	O MODELO CRISP-DM.....	18
3.2	ENTENDIMENTO DO NEGÓCIO	20
3.2.1	Controle Interno.....	20
3.2.2	O Sistema de Controle Interno do Poder Executivo Federal.....	21
3.2.3	A Controladoria-Geral da União.....	23
3.2.4	Atuação da CGU	26
3.2.5	Metas do Trabalho de Mineração	28
3.3	ENTENDIMENTO DOS DADOS	29
3.4	PREPARAÇÃO DOS DADOS	30
3.5	MODELAGEM DE DADOS	31
3.6	TEXT CLASSIFICATION E MACHINE LEARNING	32
3.6.1	Text Classification.....	32
3.6.2	Regressão Logística.....	33
3.6.3	Naive Bayes.....	34
3.6.4	Suporte Vector Machine.....	34
3.6.5	Redes Neurais.....	35
3.6.6	Avaliação dos Modelos	37
3.6.7	Métricas de Classificação	38
3.7	REFERENCIAL TEÓRICO	40
3.7.1	Recomendações da CGU	40
3.7.2	<i>Text Classification</i>.....	41

3.8	RESULTADOS.....	43
3.8.1	Validação Cruzada com os Dados de Treinamento.....	43
3.8.2	Regressão Logística.....	44
3.8.3	Multinomial Naive Bayes	45
3.8.4	Support Vector Machine.....	46
3.8.5	Redes Neurais.....	47
3.8.6	Comparação entre os Modelos.....	48
4	CONCLUSÃO.....	49
5	REFERÊNCIAS.....	50

1 INTRODUÇÃO

As atividades de auditoria desenvolvidas pela Controladoria-Geral da União – CGU apresentam um processo de trabalho que envolve, a grosso modo, as seguintes etapas: planejamento, execução e monitoramento. Na execução, a atividade final é a emissão do relatório de auditoria, com achados de auditoria e recomendações para tratamento das fragilidades identificadas.

Com relação ao monitoramento, ele consiste no acompanhamento das recomendações emitidas, com vistas a avaliar as providências adotadas pelo órgão auditado.

No momento atual, a CGU cada vez mais reserva importância à etapa de monitoramento em suas auditorias, uma vez que é nessa fase onde os resultados da auditoria se tornam visíveis, com melhoria da gestão do ente auditado. Nesse sentido, existe a preocupação de se elaborar recomendações que sejam passíveis de atendimento pelo gestor.

No processo de monitoramento, as recomendações podem se encontrar em uma das situações descritas no Quadro 1.

Quadro 1: Situações em que uma recomendação emitida pela CGU pode estar no processo de monitoramento.

Situação	Descrição
Registrada	Recomendação inserida via Sistema Monitor, mas ainda sem informação de encaminhamento ao gestor. Pode ser excluída ou alterada até o momento de registro do envio.
Atendida	Recomendação que teve posicionamento da CGU favorável à manifestação do gestor.
Monitorando	Recomendação encaminhada ao gestor e pendente de providências.
Cancelada	Recomendação cancelada pela CGU.
Consolidada	Recomendações de uma mesma Unidade Auditada, de conteúdos idênticos, agrupadas com a finalidade de racionalizar os trabalhos de monitoramento.
Não Monitorada	Recomendações que deixaram de ser monitoradas pela CGU, apesar de continuarem passíveis de implementação pelo gestor.

Fonte: CGU (2018).

Assim, o término ideal do monitoramento acontece com o atendimento da recomendação, sendo as demais situações não desejáveis do ponto de vista de eficiência do processo.

Dessa forma, o presente trabalho busca atender um dos resultados da CGU que é a o aumento da eficiência do Estado e da qualidade da entrega à sociedade¹, por meio do atendimento das recomendações formuladas na auditoria.

Para tanto, foi desenvolvido um modelo que receberá como entrada o texto de determinada recomendação escrita pela equipe de auditoria e dará, como saída, um indicador se essa recomendação será ou não atendida pelo gestor, fazendo uma espécie de controle de qualidade do trabalho desenvolvido pelo auditor.

Esse modelo será utilizado para responder a seguinte pergunta:

- É possível prever o atendimento das recomendações emitidas nos relatórios de auditoria produzidos pela CGU?

¹ Segundo o Mapa Estratégico da Controladoria-Geral da União 2020-2023, disponível em <https://www.gov.br/cgu/pt-br/aceso-a-informacao/governanca/planejamento-estrategico/arquivos/mapa-estrategico-cgu.pdf> (acesso em 1º/03/2020).

2 OBJETIVOS

2.1 Objetivo Geral

Desenvolver um classificador de texto das recomendações emitidas pela CGU, com vistas a identificar a viabilidade de atendimento de uma recomendação pelo gestor público federal.

2.2 Objetivo Específico

Escolher o algoritmo de inteligência artificial mais adequado para realizar *text classification* da base de recomendações da CGU, obtendo uma taxa de acerto nos testes acima de 75%.

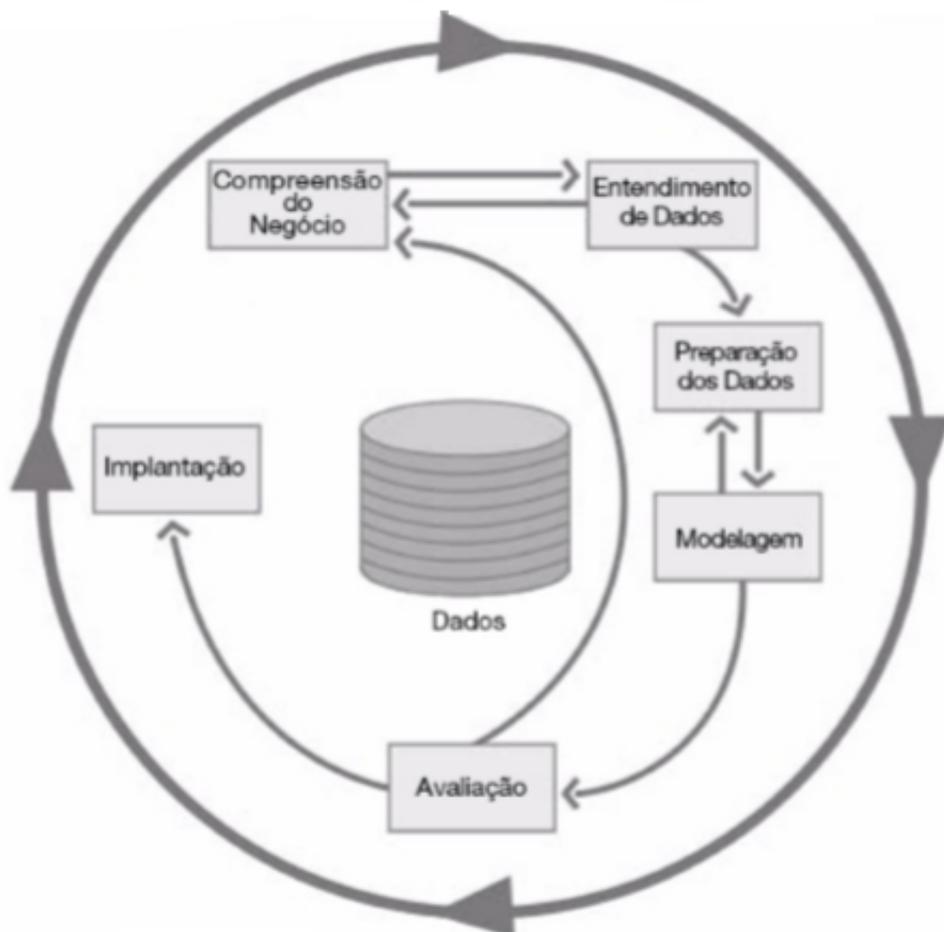
3 DESENVOLVIMENTO

3.1 O MODELO CRISP-DM

A metodologia utilizada no presente trabalho foi baseada no modelo de processo de mineração CROSS-Industry Standard Process for Data Mining – CRISP-DM, criado na década de 1990 com o objetivo de ser um modelo para as atividades de mineração de dados em bases de dados (CHAPMAN *et al.*, 2000).

As etapas do CRISP-DM estão apresentadas na Figura 1.

Figura 1: Etapas do Modelo CRISP-DM.



Fonte: Chapman *et al.* (2000).

Observa-se que o modelo se inicia pela compreensão do negócio e finaliza com a implantação de um modelo. Essas etapas podem se repetir, em um processo iterativo, e o fluxo do processo pode avançar ou voltar para a etapa anterior quantas vezes forem necessário.

As etapas do modelo estão detalhadas a seguir:

1. **Entendimento do negócio:** nesta fase, são identificados os objetivos e requisitos do projeto sob a perspectiva do negócio para então converter estas metas em um projeto de mineração de dados. Deve-se entender o background do negócio, os critérios de sucesso do projeto, o orçamento, as ferramentas disponíveis, a terminologia do contexto, os riscos e delinear um plano preliminar para atingir os objetivos identificados.
2. **Entendimento dos dados:** aqui é realizada a coleta e a familiarização com os dados. Com isso, é possível identificar a formatação e também os possíveis problemas de qualidade do dado. Já nesta fase, pode-se obter algumas conclusões sobre o problema e criar hipóteses.
3. **Preparação dos dados:** esta fase inclui todas as atividades necessárias para a construção do conjunto final de dados que será utilizado na fase seguinte de modelagem. Comumente, essa fase inclui a seleção de tabelas, registros e atributos; a transformação de dados para ficarem no formato adequado para as ferramentas que serão utilizadas; a derivação de novos atributos; e a reordenação de atributos.
4. **Modelagem de dados:** nesta etapa, as técnicas de modelagem são de fato aplicadas. Podem ser utilizadas várias abordagens diferentes e talvez seja necessário voltar à fase de preparação de dados para adequá-los às exigências das técnicas empregadas.
5. **Avaliação:** deve-se verificar se o modelo obtido atende às necessidades do negócio. Deve-se também avaliar se é necessário repetir alguma atividade.
6. **Implantação:** utilização do modelo obtido, seja para geração de relatórios que deem suporte à tomada de decisões, seja na forma de um programa que torne a mineração um processo repetível. É importante também delinear planos de monitoramento e manutenção do modelo.

3.2 ENTENDIMENTO DO NEGÓCIO

Nessa etapa, foram levantadas informações sobre o papel da CGU e o monitoramento das recomendações emitidas nos relatórios de auditoria, com o objetivo de entender o processo e a importância do modelo desenvolvido no presente trabalho.

3.2.1 Controle Interno

A CGU é o órgão central do Sistema de Controle Interno – SCI do Poder Executivo Federal, sendo importante entender do que se trata o conceito de controle interno.

O controle interno, é aquele realizado pelos órgãos e agentes da Administração, ou seja, pelo próprio Poder sobre seus atos. Os objetivos centrais do controle Interno são a prevenção e a correção de erros ou desvios no âmbito de cada Poder ou entidade da Administração Pública, devendo ser prioritariamente de caráter preventivo e estar voltado para a correção de eventuais desvios em relação aos padrões estabelecidos (PETER & MACHADO, 2014).

Nesse sentido, conforme Davila & Oliveira (2002):

O controle interno não é um evento ou circunstância, mas uma série de ações que permeiam as atividades de uma organização, ações essas inerentes ao estilo adotado pela gerência da organização na condução de seus negócios ou atividades.

Além do caráter corretivo, o controle interno deve atestar a conformidade dos atos de gestão e se preocupar com o alcance de metas, como pode ser vista na definição de controles internos presente na Instrução Normativa TCU nº 63, de 1º de setembro de 2010:

(...) controles internos: conjunto de atividades, planos, métodos, indicadores e procedimentos interligados, utilizado com vistas a assegurar a conformidade dos atos de gestão e a concorrer para que os objetivos e metas estabelecidos para as unidades jurisdicionadas sejam alcançados;

Para tanto, entende-se que as atividades de controle não podem estar desassociadas do planejamento institucional, onde serão encontrados os padrões e metas que serão utilizados no processo de avaliação e controle.

Outros componentes do controle interno, como a orientação aos gestores e o gerenciamento de riscos, podem ser vistos na definição de controle interno do Manual de

Integridade da CGU (CGU, 2013), segundo a qual o controle interno é um processo amplo, permanente e contínuo, visando ao monitoramento permanente e à fiscalização das atividades planejadas e executadas, à correção de irregularidades e à orientação dos gestores, à avaliação e o gerenciamento de riscos.

Registra-se, por fim, que a definição de controle interno do Committee of Sponsoring Organizations of the Treadway Commission – COSO é bastante afeta ao ateste de conformidade:

Controle interno é um processo conduzido pela estrutura de governança, administração e outros profissionais da entidade, e desenvolvido para proporcionar segurança razoável com respeito à realização dos objetivos relacionados a operações, divulgação e conformidade. (COSO, 2013).

Assim, por todo o exposto, depreende-se que o controle interno é o conjunto de atividades responsáveis por prevenir irregularidades, atestar a conformidade da gestão e verificar o atendimento de metas da Administração Pública. No âmbito do Poder Executivo, ele é realizado por meio do SCI, sistema que será detalhado na seção a seguir.

3.2.2 O Sistema de Controle Interno do Poder Executivo Federal

O exercício do controle interno no Poder Executivo Federal é realizado por meio do SCI, criado por meio da Lei nº 10.180, de 06 de fevereiro de 2001.

O SCI tem, como origem, o artigo 76 da Lei nº 4.320, de 17 de março de 1964. À época, o controle interno do Poder Executivo tinha como objetivo exercer o controle da execução orçamentária, compreendendo:

- I - a legalidade dos atos de que resultem a arrecadação da receita ou a realização da despesa, o nascimento ou a extinção de direitos e obrigações;
- II - a fidelidade funcional dos agentes da administração, responsáveis por bens e valores públicos;
- III - o cumprimento do programa de trabalho expresso em termos monetários e em termos de realização de obras e prestação de serviços. (art. 75 da Lei nº 4.320/1964)

Modernamente, o art. 74 da Constituição Federal de 1988 definiu as seguintes finalidades para o controle interno:

- I - avaliar o cumprimento das metas previstas no plano plurianual, a execução dos programas de governo e dos orçamentos da União;
- II - comprovar a legalidade e avaliar os resultados, quanto à eficácia e eficiência, da gestão orçamentária, financeira e patrimonial nos órgãos e entidades da administração federal, bem como da aplicação de recursos públicos por entidades de direito privado;
- III - exercer o controle das operações de crédito, avais e garantias, bem como dos direitos e haveres da União;
- IV - apoiar o controle externo no exercício de sua missão institucional.

Posteriormente, a Lei de Responsabilidade Fiscal, Lei Complementar nº 101, de 04 de maio de 2000, no art. 59, atribuiu a esse sistema a fiscalização dos seguintes temas: o atingimento das metas previstas nas leis de diretrizes orçamentárias; limites e condições para a realização de operações de créditos e a inscrição em Restos a Pagar; as medidas adotadas para o retorno da despesa total com pessoal ao respectivo limite, quando ultrapassado; da mesma forma, as providências adotadas para que as dívidas consolidada e mobiliária retornem aos respectivos limites; e a destinação dos recursos obtidos com a alienação de ativos.

Por fim, a Lei nº 10.180, de 06 de fevereiro de 2001, alterada pela Lei nº 13.464, de 10 de julho de 2017, determinou as competências do SCI:

Art. 24. Compete aos órgãos e às unidades do Sistema de Controle Interno do Poder Executivo Federal:

- I - avaliar o cumprimento das metas estabelecidas no plano plurianual;
- II - fiscalizar e avaliar a execução dos programas de governo, inclusive ações descentralizadas realizadas à conta de recursos oriundos dos Orçamentos da União, quanto ao nível de execução das metas e objetivos estabelecidos e à qualidade do gerenciamento;
- III - avaliar a execução dos orçamentos da União;
- IV - exercer o controle das operações de crédito, avais, garantias, direitos e haveres da União;
- V - fornecer informações sobre a situação físico-financeira dos projetos e das atividades constantes dos orçamentos da União;
- VI - realizar auditoria sobre a gestão dos recursos públicos federais sob a responsabilidade de órgãos e entidades públicos e privados;
- VII - apurar os atos ou fatos inquinados de ilegais ou irregulares, praticados por agentes públicos ou privados, na utilização de recursos públicos federais e, quando for o caso, comunicar à unidade responsável pela contabilidade para as providências cabíveis;
- VIII - realizar auditorias nos sistemas contábil, financeiro, orçamentário, de pessoal e demais sistemas administrativos e operacionais;
- IX - avaliar o desempenho da auditoria interna das entidades da administração indireta federal;
- X - elaborar a Prestação de Contas Anual do Presidente da República a ser encaminhada ao Congresso Nacional, nos termos do art. 84, inciso XXIV, da Constituição Federal;
- XI - criar condições para o exercício do controle social sobre os programas contemplados com recursos oriundos dos orçamentos da União.

Ou seja, percebe-se uma evolução das atribuições do SCI. Na sua criação, ele era essencialmente voltado para a fiscalização das despesas da União, com vistas a atestar sua regularidade. Hoje, o sistema estendeu suas competências para a avaliação dos programas de governo e do cumprimento de metas governamentais, atividades que agregam, de fato, valor à gestão.

Em resumo, o SCI tem base constitucional, atuação sobre todos os órgãos e entidades do Poder Executivo Federal e competência para fiscalizar a execução de recursos públicos descentralizados provenientes do orçamento da União para organizações privadas e para outros entes federativos (MUNIZ, 2017).

Ressalta-se que o controle no modelo brasileiro é bipartido, em controle interno e externo, conforme art. 71 da Constituição Federal de 1988, que estabelece que o Controle Externo será realizado pelo Congresso Nacional, via Tribunal de Contas da União – TCU e o Controle Interno pelo sistema de controle interno de cada poder, de forma complementar ao controle externo. (AZEVEDO, 2016)

Ou seja, as atividades de controle interno e controle externo não são executadas totalmente separadas, mas sim, atuam em colaboração uma com a outra. Muitas das atividades de controle iniciam-se nos sistemas de controle interno e são concluídas no controle externo, a quem cabe julgar as contas dos administradores públicos e demais responsáveis por dinheiros, bens e valores públicos. (MARX, 2015)

Nesse sentido, o SCI tem um papel importante no auxílio à gestão, procedendo à apreciação dos atos da Administração, dando ciência a autoridade principal e emitindo um parecer que são objetos da prestação de contas anual. (SILVA, 2016)

No que concerne à estrutura do SCI, ele é formado pela Secretaria Federal de Controle Interno – SFC, órgão integrante da CGU; e por órgãos setoriais, responsáveis pelo controle interno do ministério das Relações Exteriores, do Ministério da Defesa, da Advocacia-Geral da União e da Casa Civil.

3.2.3 A Controladoria-Geral da União

A CGU foi criada pela Lei nº 10.683, de 28 de maio de 2003, e apresenta o status de Ministério, conforme art. 19 da Lei nº 13.844, de 18 de junho de 2019. De acordo com o art. 51 desta mesma lei, constituem áreas de competência da CGU:

- I - providências necessárias à defesa do patrimônio público, ao controle interno, à auditoria pública, à correição, à prevenção e ao combate à corrupção, às atividades de ouvidoria e ao incremento da transparência da gestão no âmbito da administração pública federal;
- II - decisão preliminar acerca de representações ou denúncias fundamentadas recebidas e indicação das providências cabíveis;
- III - instauração de procedimentos e processos administrativos a seu cargo, com a constituição de comissões, e requisição de instauração daqueles injustificadamente retardados pela autoridade responsável;
- IV - acompanhamento de procedimentos e processos administrativos em curso em órgãos ou entidades da administração pública federal;
- V - realização de inspeções e avocação de procedimentos e processos em curso na administração pública federal, para exame de sua regularidade, e proposição de providências ou correção de falhas;
- VI - efetivação ou promoção da declaração da nulidade de procedimento ou processo administrativo em curso ou já julgado por qualquer autoridade do Poder Executivo federal e, se for o caso, da apuração imediata e regular dos fatos envolvidos nos autos e na nulidade declarada;
- VII - requisição de dados, de informações e de documentos relativos a procedimentos e processos administrativos já arquivados por autoridade da administração pública federal;
- VIII - requisição a órgão ou a entidade da administração pública federal de informações e de documentos necessários a seus trabalhos ou a suas atividades;
- IX - requisição a órgãos ou a entidades da administração pública federal de servidores ou de empregados necessários à constituição de comissões, inclusive das referidas no inciso III do caput deste artigo, e de qualquer servidor ou empregado indispensável à instrução de processo ou procedimento;
- X - proposição de medidas legislativas ou administrativas e sugestão de ações para evitar a repetição de irregularidades constatadas;
- XI - recebimento de reclamações relativas à prestação de serviços públicos em geral e à apuração do exercício negligente de cargo, emprego ou função na administração pública federal, quando não houver disposição legal que atribua essas competências específicas a outros órgãos;
- XII - coordenação e gestão do Sistema de Controle Interno do Poder Executivo federal; e
- XIII - execução das atividades de controladoria no âmbito da administração pública federal.

A importância da Secretaria Federal de Controle Interno nesse contexto se dá em função de ser por meio dela que a CGU exerce seu papel de Órgão Central do SCI.

Segundo Olivieri (2008), as principais funções da SFC, no exercício da competência de órgão central do SCI, são gerenciar e normatizar sobre atividades de controle interno do

governo federal, com atribuições delineadas no art. 70 da Constituição Federal de 1988. Para a autora, o controle interno governamental consiste no conjunto de atividades de auditoria e fiscalização que busca, além da garantia de conformidade legal dos atos da administração pública, prover os gestores com instrumentos de monitoramento quanto à eficiência, economicidade e eficácia de suas ações, objetivando mitigar perdas, aprimorar a gestão pública e garantir a prestação pública de contas

Cabe mencionar que a CGU – além do Sistema de Controle Interno – também é órgão central dos sistemas de Correição e Ouvidoria do Poder Executivo Federal e está estruturada em quatro grandes áreas com status de Secretarias, a saber: Secretaria Federal de Controle Interno; Corregedoria-Geral da União; Ouvidoria Geral da União; e Secretaria de Transparência e Prevenção à Corrupção. (DIAS, 2018).

A partir da vigência do Decreto nº 9.681, de 3 de janeiro de 2019, a CGU ganhou mais uma secretaria: a Secretaria de Combate à Corrupção, especializada em acordos de leniência, operações especiais e pesquisa e informações estratégicas.

Nos últimos anos a CGU tem sido reconhecida pela população e dentro da Administração Pública como um órgão do governo federal responsável por ações de combate à corrupção. Nesse sentido, Avritzer (2016) classifica a CGU, o TCU e a Polícia Federal como instituições de controle criadas pelo Estado brasileiro a partir de 1988 e que têm desempenhado um importante papel na limitação da corrupção.

A própria Constituição Federal de 1988, assim como os demais normativos e os Planos Estratégicos do Órgão, indicam que o combate à corrupção é um dos efeitos da atuação da CGU, mas não necessariamente representa seu principal objetivo como órgão de Estado, principalmente quando se trata da função de controle interno. (ARANTES, 2016)

Pelo contrário, a atuação da CGU como órgão de controle interno tem a principal finalidade de promover o aperfeiçoamento e a transparência da gestão pública por meio da avaliação e do controle das políticas públicas e da qualidade dos gastos. Para atingir esse objetivo, é necessário que o órgão de controle exerça o papel de assessoramento da gestão, provendo subsídios para a sua melhoria constante e para o fortalecimento do ambiente interno das organizações públicas. Essa forma de atuação se contrapõe, em certos momentos, à atuação eminentemente punitiva que o órgão adota diante de desvios e malfeitos com recursos públicos. Dessa forma, a consolidação do controle interno como atividade essencial para o funcionamento

da administração pública exige o debate acerca da dualidade das ações de controle, que devem incorrer em atos punitivos quando necessários, mas que devem ser úteis para o aprimoramento da gestão. (ARANTES, 2016)

Marx (2015) também reconhece essa dualidade na atuação da CGU: como órgão de combate à corrupção e como órgão integrante do ciclo de gestão que atua no aprimoramento das políticas públicas.

3.2.4 Atuação da CGU

A CGU desenvolve dois tipos de atividade: as atividades de avaliação, com um viés mais de auditoria, e as de consultoria, prestando apoio no desenvolvimento de políticas públicas e na melhoria da eficiência do Poder Executivo Federal.

O Controle Interno do Poder Executivo Federal atua sobre áreas tão díspares como educação, saúde e segurança, cada uma com seus conhecimentos específicos. Não à toa, os concursos públicos para o cargo de nível superior da CGU (Auditor Federal de Finanças e Controle) não cobram uma graduação específica como contabilidade ou administração. São contratados auditores de todas as formações possíveis, desde um biólogo até um engenheiro eletrônico, por exemplo.

Diante do quantitativo extenso de campos onde a CGU pode atuar, faz-se necessário escolher as áreas que receberão mais atenção pelos auditores. Nesse contexto, a priorização das ações de controle desenvolvidas pela CGU vem passando por um processo de mudança. Antes, as ações eram escolhidas com base nos critérios de materialidade, criticidade e relevância. Atualmente, o foco é a realização de auditorias baseadas em riscos, sendo primordial o conhecimento a fundo das Unidades Auditadas.

Após a definição do escopo de atuação da CGU, são realizadas as auditorias, que se iniciam com uma etapa de planejamento, onde são levantados os pontos críticos do processo auditado e são elaboradas questões de auditoria que serão respondidas ao longo dos trabalhos.

A próxima etapa é o trabalho de campo, onde a CGU irá coletar as evidências que suportarão as avaliações relacionadas a cada questão de auditoria.

Não necessariamente o período de campo ocorrerá fisicamente nas instalações da Unidade Auditada, podendo as análises serem realizadas quase que completamente na mesa do auditor, como no caso de auditorias baseadas em cruzamentos de bases de dados.

Com base nas evidências coletadas e produzidas, a equipe de auditoria elabora o Relatório de Auditoria.

O Relatório é o documento que reflete os resultados dos exames efetuados pelo SCI, os quais são levados ao conhecimento das autoridades competentes, com as finalidades de fornecer dados para a tomada de decisão, para atendimento das recomendações sob sua responsabilidade e para correção de erros detectados (DIAS, 2018).

Um Relatório de Auditoria tem, no seu cerne, uma sequência de achados de auditorias, os quais estão relacionados às questões de auditoria e às recomendações.

As recomendações consistem em providências que devem ser adotadas para correção de falhas e irregularidades, com a principal finalidade de contribuir para o aperfeiçoamento da gestão pública. Elas podem ter como objetivo a implantação de controles, a mitigação de riscos, a apuração de responsabilidade administrativa, a reposição de valores ao Erário e a atuação junto a outros órgãos.

Assume-se que as recomendações destinadas exclusivamente ao aprimoramento da gestão são aquelas que melhor representam o papel de assessoramento do controle interno, enquanto que as recomendações voltadas a responsabilização dos agentes públicos representam o extremo oposto, onde a atuação é eminentemente punitiva. (ARANTES, 2016)

A definição do conjunto de recomendações de um Relatório de Auditoria é feita pela equipe de auditoria em conjunto com os gestores, em um encontro denominado Reunião de Busca Conjunta de Soluções.

De acordo com o Manual de Elaboração de Relatórios do Controle Interno (CGU, 2014), as recomendações devem: concentrar-se no propósito de eliminar, mitigar ou reduzir as causas das falhas encontradas; estar de acordo com o princípio da racionalização administrativa e da economia processual; ser claras quanto às providências a serem adotadas pelo gestor, evitando expressões genéricas; e ser tempestivas e exequíveis.

Esse arcabouço de critérios e preceitos direcionados ao processo de elaboração de uma recomendação tem dois objetivos: garantir que as recomendações agreguem valor à gestão e proporcionar seu atendimento após a adoção de providências por parte da unidade auditada.

Ressaltando que o mais relevante não é meramente o cumprimento da recomendação, mas sim a efetiva resolução do problema identificado. (AZEVEDO, 2016)

Nesse contexto, segundo Madeira (2013), para que o relatório produza resultados, torna-se importante a *follow-up* das recomendações nele consignadas, ou seja, o monitoramento das recomendações.

Nessa etapa, são realizadas atividades destinadas a verificar a adoção por parte dos gestores da unidade examinada de providências destinadas a sanar as causas dos fatos relatados, a recomposição de valores ou a responsabilização de agentes, quando cabível. O monitoramento de uma recomendação se encerra com o seu pleno atendimento ou com o seu cancelamento, seja pela perda do objeto ou pela inconveniência de se manter o acompanhamento. (ARANTES, 2016)

O processo de monitoramento das recomendações ocorre por meio do Plano de Providências Permanente – PPP, assim definido no Manual do Sistema Monitor (CGU, 2014):

(...) documento elaborado por unidade integrante da administração federal direta ou indireta, com base nas recomendações do Controle Interno, onde constam atualizadas as providências que os gestores devem adotar para regularizar ou sanear as falhas apontadas em ações de controle.

Até 2015, o PPP era uma planilha encaminhada semestralmente aos órgãos auditados onde cada linha do documento era uma recomendação em monitoramento. Por meio de ofícios trocados entre a CGU e os demais órgãos do Poder Executivo Federal, eram informadas as providências adotadas para o cumprimento das recomendações.

Esse processo burocrático e centrado no papel foi substituído pelo Sistema Monitor, lançado em 2015 pela CGU. Em 2019, ele foi substituído pelo e-Aud.

O e-Aud é o sistema de gestão da Atividade de Auditoria Interna Governamental, integrando todo o processo de auditoria, desde o mapeamento do universo auditável e o desenvolvimento do plano anual de auditoria, até a operacionalização dos trabalhos individuais, contemplando as fases de planejamento, execução, comunicação dos resultados e monitoramento das recomendações emitidas, entre outras funcionalidades.

3.2.5 Metas do Trabalho de Mineração

Tendo em vista que o principal critério de sucesso de um trabalho de auditoria realizada pela CGU é o atendimento das recomendações emitidas no relatório, propõe-se a elaboração de um preditor do atendimento dessas recomendações.

O foco deve ser a orientação sobre qual a melhor forma de se escrever uma recomendação de auditoria para seu atendimento. Esse é o objetivo do presente trabalho, por meio do desenvolvimento de um modelo de *text classification* para prever o comportamento do gestor federal frente a uma recomendação.

Esse preditor poderá ser utilizado no momento da escrita da recomendação de auditoria, informando ao auditor uma previsão do estado final da recomendação, se ela será atendida ou não. A partir dessa informação, o auditor poderá reavaliar a recomendação e verificar junto às orientações estabelecidas pela CGU se a referida recomendação segue todos os padrões de escrita estipulados pela organização.

Para garantir o sucesso da implantação do modelo da CGU, definiu-se a meta de 75% de precisão do modelo. Com isso, espera-se que o percentual de atendimento das recomendações emitidas pela CGU ultrapasse os atuais 69,6% de atendimento², alcançando uma linha de base de 75% de atendimento das recomendações.

3.3 ENTENDIMENTO DOS DADOS

Os dados referentes às recomendações formuladas pela CGU são armazenados em dois sistemas: Novo Ativa (sistema para registro das atividades de auditoria) e Monitor (registro do monitoramento das recomendações). Ambos os sistemas possuem bases de dados em servidor SQL Server mantido pela CGU.

O Novo Ativa era responsável por armazenar as informações de monitoramento das recomendações até a implantação do sistema Monitor em 2015. Assim, a tabela com a descrição da situação da recomendação está armazenada no banco de dados do Novo Ativa e foi utilizada no presente trabalho.

Com relação ao banco de dados do sistema Monitor, foi utilizada a tabela base de recomendações, onde estão registrados o texto da recomendação e a situação da recomendação. Nessa tabela, havia 179.824 recomendações registradas na data de 19/12/2019, quando ocorreu a extração dos dados utilizados no presente trabalho.

² Define-se como o percentual de atendimento das recomendações a relação entre as recomendações atendidas e as recomendações canceladas ou não monitoradas, desprezando as recomendações consolidadas e as que se apresentam em estágios intermediários do processo.

Foram aplicados filtros nessa base, sendo selecionadas as seguintes recomendações para modelagem: recomendações enviadas aos gestores a partir de 2009, recomendações nas situações atendida, cancelada e não monitorada.

As demais situações (registrada, monitorando e consolidada) foram excluídas do modelo por se tratarem de situações intermediárias, onde ainda existe chance da recomendação ser atendida (situações “registrada” e “monitorando”), ou porque a recomendação foi agrupada em outra recomendação semelhante (situação “consolidada”).

3.4 PREPARAÇÃO DOS DADOS

Inicialmente, foram excluídas as recomendações com textos idênticos e aquelas consideradas erros na base (textos sem sentido, textos que não representam recomendações e textos de recomendações de teste que ainda estavam registrados na base), a partir da avaliação de especialista no tema.

Ainda, retirou-se os *outliers* da base, com base no tamanho do texto da recomendação (ou seja, quantidade de caracteres). Para recomendações muito pequenas existe um risco relevante de serem erros na base. Recomendações muito grandes indicam que houve erro na elaboração da recomendação de auditoria, uma vez que o excesso de detalhamento da solução na recomendação pode engessar o desenvolvimento de providências pelo gestor para atendimento da recomendação. Assim, considerou-se como *outliers* aquelas recomendações cuja quantidade de caracteres era dois desvios padrão maior ou menor do que a média.

Após essas exclusões, a base de dados final que foi utilizada para modelagem ficou com 65.971 registros, sendo 45.931 recomendações atendidas (69,6% do total), 12.608 recomendações canceladas (19,1% do total) e 7.432 recomendações não monitoradas (11,3% do total).

Na sequência, o texto das 65.971 recomendações passaram pelo seguinte tratamento:

1. Selecionou-se apenas o texto da recomendação e a respectiva situação da recomendação, a qual foi ajustada para 0 = atendida, 1 = demais situações.
2. Essas informações foram extraídas para um arquivo csv e importado para um notebook Jupyter, onde o campo com o texto da recomendação recebeu os seguintes ajustes: as letras foram convertidas para minúsculo; os sinais de pontuação foram excluídos; e as *stopwords* foram eliminadas.

3. Foi utilizada a técnica BOW para construção de um vetor para representação dos textos das recomendações, para fins de treinamento e teste do modelo.
4. A quantidade de palavras a ser guardada no BOW foi ajustada para 9.000, a partir de testes realizados com RL³ que demonstraram que esse valor era o mais adequado para a base de recomendações.
5. Para escolha das 9.000 palavras mais representativas dos textos das recomendações escolhidas para fazerem parte do BOW, foi utilizada a medida estatística *Term Frequency - Inverse Document Frequency* – TF-IDF.
6. O conjunto de dados foi dividido em dois subconjuntos: um subconjunto de treino (52.776 recomendações – 80% do total) e um de testes (13.195 recomendações), que será utilizado apenas na etapa de validação.
7. Tendo em vista que o conjunto de dados estava desbalanceado, com 69,6% das recomendações classificadas como atendidas, foi utilizada a técnica de *resample (oversampling)* para igualar a quantidade de recomendações atendidas com a quantidade de recomendações não atendidas na base de treino. Após o balanceamento, essa base passou a ter 73.744 registros.

3.5 MODELAGEM DE DADOS

A ideia de modelagem, foi a construção de um classificador de recomendações que teve, como entrada, o texto de uma recomendação e, como saída, o potencial dela se tornar uma recomendação atendida.

Na segunda iteração, utilizou-se as seguintes técnicas de *Machine Learning* – ML para escolha do modelo: Regressão Logística – RL, Multinomial Naive Bayes – NB, *Support Vector Machines* – SVM e Redes Neurais Perceptron Multi-Camadas (ou Multilayer Perceptron – MLP).

A etapa de avaliação da modelagem será apresentada na seção 3.8, como resultado do presente trabalho. A seção a seguir detalha as técnicas de ML utilizadas na execução do trabalho.

³ Os testes envolveram rodar uma base preliminar dos dados em uma RL com diversos tamanhos de BOW, sendo que o melhor resultado foi obtido com 9.000 palavras.

3.6 TEXT CLASSIFICATION E MACHINE LEARNING

De uma maneira geral, aprendizagem de máquina (ou ML) é a área de estudo que se preocupa em construir soluções computacionais que automaticamente se aperfeiçoam com suas experiências (MITCHELL, 1997).

Pode-se dizer que o ML é uma vertente da Inteligência Artificial (do inglês *Artificial Intelligence* – AI), o qual possui um corpo de conhecimento específico ou um conjunto de técnicas para que a máquina possa aprender com os dados. A AI é um campo de estudo muito mais abrangente que engloba os campos de visão computacional, processamento de linguagem natural, robótica, entre outros (BRINK *et al.*, 2016).

Os algoritmos de ML dependem da aplicação em questão e do conhecimento *a priori* que se tem do domínio ao qual a aplicação faz parte.

Entre esses algoritmos, existem aqueles relacionados à aprendizagem supervisionada, onde há um conjunto rotulado de dados de treinamento com uma série de pares entrada-saída. Esse conjunto de treinamento será processados pelo modelo, com o objetivo de se obter uma função que consiga classificar novas entradas.

Observe que a aprendizagem supervisionada é a que será utilizada no presente trabalho porque a solução para o problema em questão busca identificar de qual classe a entrada pertence. E como os dados já possuem um histórico classificado, o trabalho seria facilitado.

3.6.1 Text Classification

A classificação de textos, quando realizada por um computador e utilizando o Processamento de Linguagem Natural – PLN, consiste em detectar a categoria a qual pertence determinado texto com base nas características observadas e na probabilidade dessas características pertencerem a determinadas classes pré-definidas (RUSSEL & NORVIG, 2013).

O ponto de partida para a análise de textos em ML se dá pelo desmembramento dos dados de entrada em objetos menores (*tokens*) que possam ser analisados individualmente. Isso ocorre na etapa de pré-processamento dos dados.

Para essa etapa, duas técnicas foram empregadas no presente trabalho: tokenização e remoção de *stopwords*.

A tokenização é um processo que separa as palavras de um documento, removendo as pontuações, e substituindo os elementos não textuais, como quebra de linha e tabulação, por um espaço em branco simples (HOTHOTH *et al.*, 2005).

As *stopwords* são as palavras que pouco acrescentam à representatividade da coleção de dados, ou que sozinhas nada significam. Exemplos de *stopwords* são palavras como artigos, pronomes e advérbios. O conjunto de *stopwords* é chamado de *stoplist*. Essa eliminação de *stopwords* reduz significativamente a quantidade de termos, diminuindo o custo computacional das próximas etapas (REZENDE *et al.*, 2011).

Após o pré-processamento, os textos das recomendações foram reorganizados em vetores, utilizando BOW.

BOW é uma estratégia de representação do texto em que a sequência de palavras é desconsiderada. Dessa maneira, o texto é simplificado para um vetor de palavras distintas e a respectiva contagem de ocorrência de cada uma delas. A principal vantagem dessa estratégia é a sua simplicidade de representação computacional, sendo atrativa para a aplicação direta de algoritmos de aprendizado de máquina.

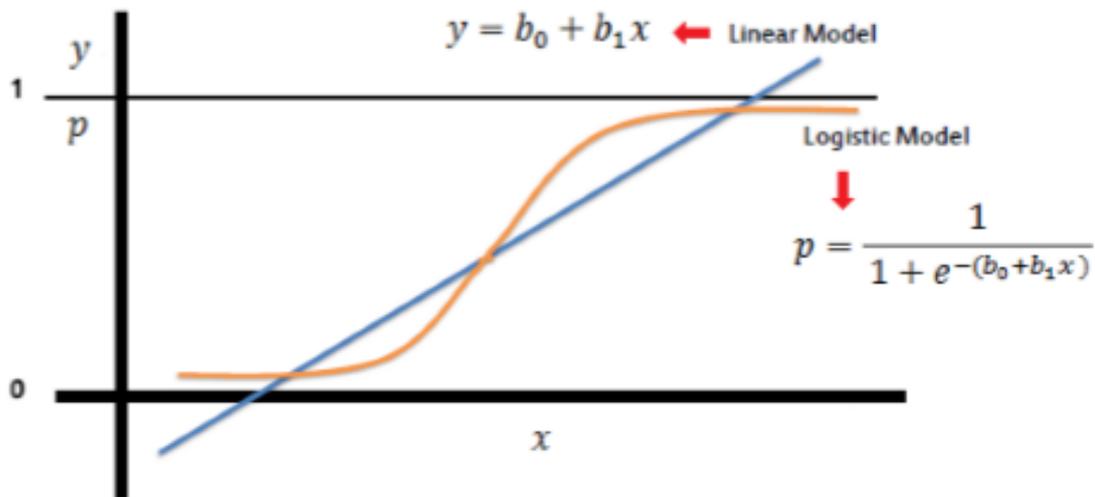
A escolha das palavras que foram incluídas no vetor se deu por TF-IDF, que é a combinação de duas técnicas estatísticas. *Term Frequency* mede o quão frequente um termo ocorre em um documento e o *Inverse Document Frequency* mede a raridade do termo para o documento. Uma medida razoável da importância de um termo pode ser obtida utilizando o dessa duas estatísticas (SALTON & BUCKLEY, 1988).

3.6.2 Regressão Logística

A RL permite a identificação probabilística de que uma série de variáveis é responsável por um resultado (SAYAD, 2017).

A RL é mais adequada em relação à regressão linear para avaliar um modelo cujo resultado é binário, como é o caso de um preditor de atendimento ou não de recomendações, e isso acontece porque a RL não permite valores fora do intervalo [0 ... 1], como pode ser visto na Figura 2.

Figura 2: Comparação entre a regressão linear e a regressão logística.



Fonte: Sayad (2017).

3.6.3 Naive Bayes

Thomas Bayes, um pastor nascido em Londres em 1701, foi quem deu os primeiros passos ao estudo probabilístico, descrevendo que a probabilidade de eventos não conectados ocorrer é menor do que a probabilidade de eventos conectados ocorrer (MLODINOW, 2009).

O Teorema de Bayes é uma fórmula matemática usada para o cálculo da probabilidade de um evento dado que outro evento já ocorreu, o que é chamado de probabilidade condicional. A partir desse teorema, é possível construir um classificador muito simples que presume a independência dos dados de entrada⁴.

Segundo Han & Kamber (2011) a classificação bayesiana é uma técnica eficaz para classificação de documentos, sendo utilizada, por isso, no presente trabalho.

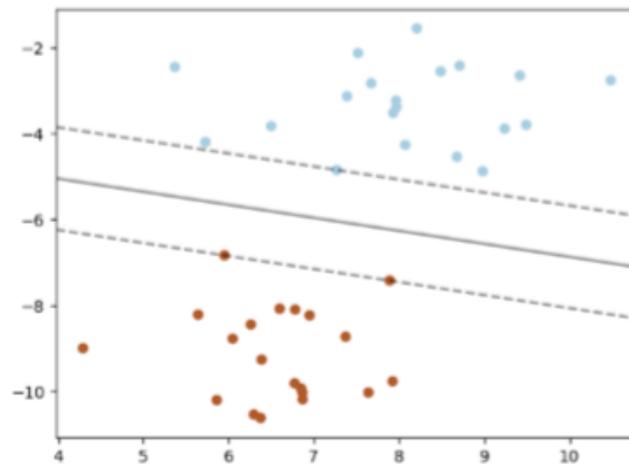
3.6.4 Suporte Vector Machine

Uma SVM recebe os vetores de entrada e faz transformações utilizando seus núcleos para que seja possível aplicar métodos lineares, mesmo em problemas não-lineares (HEARST *et al.*, 1998).

⁴ Como o classificador presume a independência entre os dados de entrada, ele é chamado de ingênuo (*naive*, em inglês).

A função de classificação do SVM é baseada em hiperplanos. O algoritmo busca otimiza-los de forma que eles se posicionem ortogonalmente às linhas que separam as classes do problema, como pode ser visto na Figura 3.

Figura 3: Funcionamento de uma SVM, mostrando a separação das classes por meio de hiperplanos.



Fonte: Pedregosa *et al.* (2011).

Havendo múltiplos hiperplanos possíveis de satisfazer, o algoritmo seleciona aquele em que haja a maior margem entre o hiperplano e os indivíduos mais próximos. A maximização desta margem reduz a chance de erros quando novas instâncias forem classificadas (BOSER *et al.*, 1992).

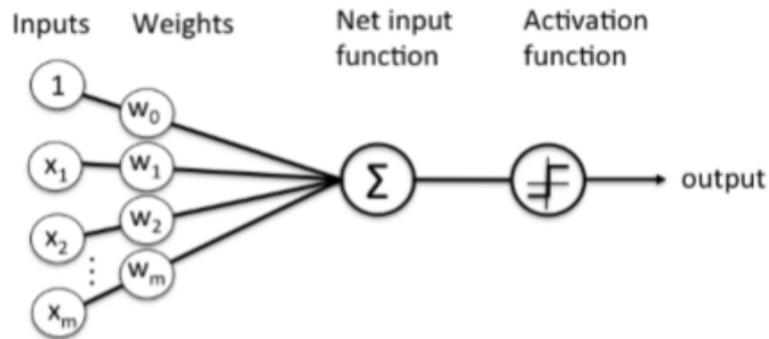
3.6.5 Redes Neurais

Um dos algoritmos desenvolvidos para realizar o ML foi baseado em um neurônio humano (GOLDBERG, 2017).

O neurônio, chamado de perceptron, possui sensores de entrada. Os sinais recebidos nesses sensores são multiplicados por pesos e se propagam para uma área de associação, onde são combinados em um único sinal por meio de uma função de ativação. Na sequência, esse sinal é propagado para as saídas do neurônio apenas se o valor dele for maior que um limiar de excitação θ (RASCHKA, 2015).

A Figura 4 apresenta esquematicamente o fluxo de passagem do sinal por um neurônio.

Figura 4: Representação esquemática de um perceptron de uma rede neural.



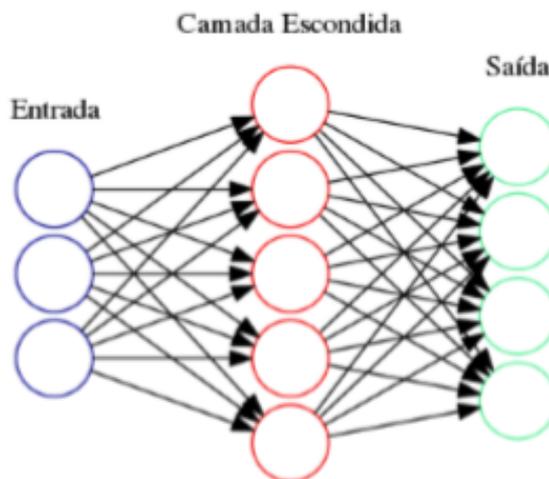
Fonte Raschka (2015).

Agrupando esses perceptrons, tem-se uma rede neural MLP. As camadas de perceptrons dessa rede podem ser classificadas em:

- Entrada: esta camada não possui funções de ativação, tendo como função a propagação dos dados de entrada para a primeira camada oculta.
- Camadas ocultas: são as camadas que possuem funções de ativação.
- Saída: é a ultima camada de uma rede neural que irá retornar o resultado do processamento realizado pelas camadas ocultas.

A Figura 5 seguir apresenta um desenho esquemático de uma rede neural MLP.

Figura 5: Representação esquemática de uma rede neural MLP.



Fonte: Santos (2017).

A MLP foi a rede neural mais utilizada em problemas de aprendizagem de máquinas durante muito tempo. Sua popularidade foi oriunda do seu baixo custo computacional e sua capacidade de distinguir dados que não são linearmente separáveis, algo que seus antecessores não conseguiam fazer (SANTOS, 2017).

O fluxo comum da execução de uma rede neural é o recebimento dos dados de entrada na primeira camada, em seguida há a propagação dos valores para uma camada oculta, após o cálculo de ativação, os dados são propagados novamente para a próxima camada, até que ao final, chega-se na camada de saída (GOLDBERG, 2017).

Esse fluxo ocorre quando se utiliza a rede neural para prever a saída de dados a partir de uma entrada. Para que a rede neural tenha sucesso nessa tarefa, é preciso treiná-la antes.

Para o treinamento dos pesos da rede é usado o *back-propagation*, ou retropropagação. Durante esse processo, os pesos dos perceptrons são ajustados tendo como base uma medida de erro (NORVIG & RUSSEL, 2013).

3.6.6 Avaliação dos Modelos

O processo de avaliação dos modelos desenvolvidos no presente trabalho consistiu em dividir os dados rotulados de entrada do modelo em dois conjuntos: um conjunto de treinamento (com 80% dos dados de entrada) e um conjunto de testes (com 20% dos dados de entrada).

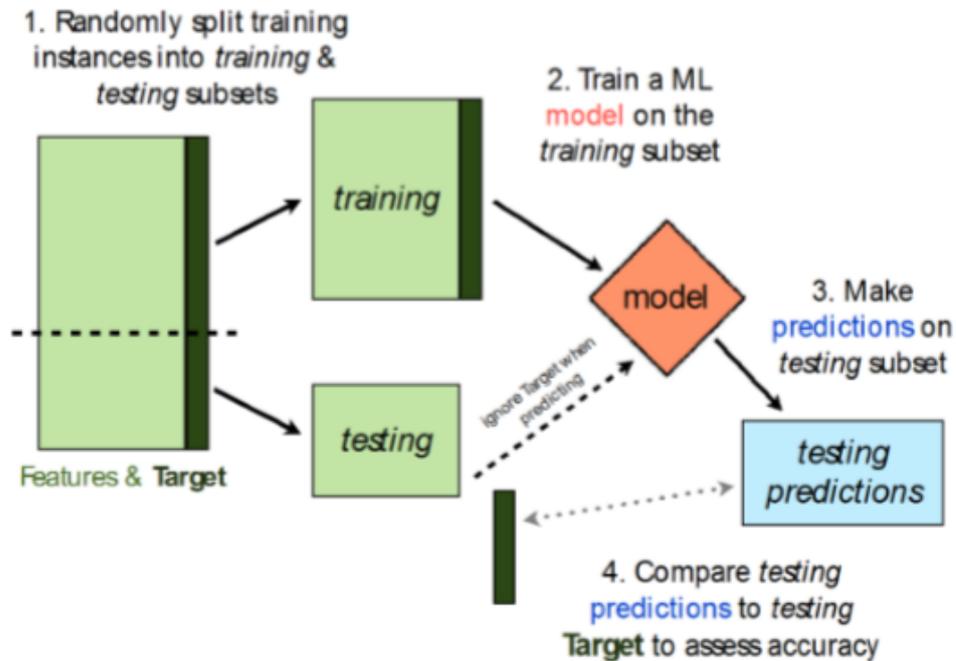
O modelo é então treinado com o conjunto de treinamento e testado com o conjunto de testes.

Para cada modelo utilizado no presente trabalhos, foram variados hiperparâmetros com vistas a encontrar o melhor conjunto de hiperparâmetros de cada um. Para escolha da melhor configuração de cada modelo, foi utilizada a estratégia de validação cruzada por meio de *k-fold* subconjuntos do conjunto de treinamento. Nessa etapa, não foi utilizado o conjunto de testes para não enviesar a escolha do melhor conjunto de hiperparâmetros.

Na etapa de testes, o modelo não recebe o rótulo dos dados de teste e faz uma predição do rótulo. Na sequência, o valor predito é comparado com o valor real do rótulo, sendo calculadas as métricas de avaliação do modelo.

A Figura 6, resume o processo de de avaliação utilizado no presente trabalho.

Figura 6: Representação esquemática do processo de avaliação dos modelos utilizados no presente trabalho.



Fonte: Brink *et al.* (2016).

3.6.7 Métricas de Classificação

Em modelos de classificação, é possível determinar com exatidão o resultado esperado em categorias. Com isso, é possível determinar o Verdadeiro Positivo e o Verdadeiro Negativo, que é quando o modelo acerta a predição. Para quando ele erra, há o Falso Positivo e o Falso Negativo (BRINK *et al.*, 2016).

A matriz de confusão apresenta a quantidade de dados de testes que foram preditos corretamente ou erroneamente, sendo representada na Figura 7.

Figura 7: Representação esquemática de uma matriz de confusão.

Matriz de Confusão	Rótulo A (A soma da coluna apresenta o total de predições tendo como resultado o Rótulo A)	Rótulo B (A soma da coluna apresenta o total de predições tendo como resultado o Rótulo B)
Rótulo A (A soma da linha apresenta o total de dados da base de testes com o rótulo A)	Verdadeiro Positivo do Rótulo A	Falso Positivo do Rótulo B
Rótulo B (A soma da linha apresenta o total de dados da base de testes com o rótulo B)	Falso Positivo do Rótulo A	Verdadeiro Positivo do Rótulo B

Fonte: Elaboração própria.

Essa matriz facilita a visualização do número de classificações corretas e do número de classificações preditas para cada classe, segundo o classificador em análise. Assim, torna-se uma ferramenta útil para analisar a qualidade de um classificador de texto (HAN & KAMBER, 2011).

Exemplificando a matriz de confusão, observe a matriz hipotética apresentada no Quadro 2.

Quadro 2: Matriz de Confusão hipotética de um modelo que avaliou 3.000 dados rotulados em duas classes.

	Rótulo A	Rótulo B
Rótulo A	1.480	25
Rótulo B	20	1.475

Fonte: Elaboração Própria.

A partir dessa matriz, podemos identificar que o modelo acertou 1.480 predições do rótulo A e errou 20 predições do rótulo A. Da mesma forma, o modelo errou 25 predições do rótulo B e acertou 1.475 predições do rótulo B.

A matriz de confusão é a base para as métricas utilizadas no presente trabalho:

- Precisão: representa a proporção de verdadeiros positivos de uma classe em relação ao total de dados preditos para essa classe. Ou seja, é a divisão entre os

verdadeiros positivos preditos pelo modelo para uma classe e o total de dados preditos para essa classe (a soma entre os verdadeiros positivos e os falsos positivos da classe). É o inverso da Taxa de Erro.

- *Recall* (ou sensibilidade, em português): é a taxa de acerto de uma classe (ou rótulo), correspondendo a divisão entre os verdadeiros positivos preditos pelo modelo para uma classe e o total de dados rotulados com a classe.
- *F1-Score*: média harmônica entre precisão e *recall*.

3.7 REFERENCIAL TEÓRICO

3.7.1 Recomendações da CGU

São poucos os trabalhos acadêmicos que apresentaram conclusões ou avaliaram as recomendações emitidas pela CGU.

Araújo (2007) realizou uma pesquisa junto aos auditores internos de Instituições Federais de Ensino – IFE, verificando que as recomendações emitidas pela CGU influenciavam a tomada de decisão dos gestores das IFE.

Hanashiro (2009), ao propor um modelo de implementação de auditoria de TI na CGU, verificou que o cenário então vigente nas auditorias de TI prejudicava a emissão de recomendações, em especial quanto à ausência de padrões e à não obrigação, pelos órgãos públicos, de seguir os referenciais de boas práticas em TI.

Azevedo (2016) desenvolveu o primeiro trabalho mais focado em avaliar a influência das recomendações dos relatórios de auditoria da CGU na melhoria da gestão, escolhendo a Secretaria Executiva do Ministério da Saúde como unidade a ser pesquisada. O autor verificou que apenas 50% das recomendações direcionadas ao órgão foram implantadas, definindo esse cenário como decorrente de um distanciamento entre os órgãos de controle e os gestores na ponta. A partir de entrevistas, consignou a visão dos gestores sobre o trabalho da CGU: as recomendações eram, em sua maioria, relacionadas a problemas de conformidades ou não eram aplicáveis, por carência de recursos humanos ou tecnológicos. Outro ponto abordado foi sobre o perfil do auditor, que se mostrou deficiente de conhecimento para auditar áreas específicas da administração pública.

Arantes (2016) realizou um estudo qualitativo acerca da relação entre o auditor da CGU e o auditado. Entre as dificuldades apontadas pelos gestores, com relação às

recomendações emitidas pela CGU, destaca-se: o foco no problema e não na solução, a existência de recomendações inócuas ou genéricas, o direcionamento da recomendação para o gestor federal incorreto, o entendimento equivocado ou insuficiente dos auditores, e a intempestividade do envio da recomendação ao gestor.

Dias (2018) identificou variáveis determinantes para o atendimento das recomendações emitidas pela CGU entre 2011 e 2015 em auditorias de obras públicas. Para o autor, equipes de auditoria heterogêneas, o maior tempo médio de experiência profissional da equipe, auditorias originadas de demandas externas e ações investigativas são as variáveis que influenciam no maior grau de atendimento das recomendações. Por outro turno, Unidades Gestoras com elevado grau de rotatividade e histórico de processos disciplinares de seus dirigentes prejudicam o atendimento de recomendações formuladas pela CGU.

Ou seja, observa-se que os trabalhos acadêmicos sobre as recomendações emitidas pela CGU buscaram identificar as causas que influenciavam o grau de atendimento das recomendações. Entretanto, não se verificou, em trabalhos anteriores,

3.7.2 Text Classification

A Tabela 7 apresenta outros classificadores de textos em português baseados em ML desenvolvidos pela academia:

Tabela 1: Compilação de trabalhos de text classification em textos em português utilizando técnicas de ML. Resultados ordenados pela precisão média do melhor modelo encontrado.

TRABALHO	TEMA	TÉCNICA	RESULTADO
Lima & Pinto (2016)	Textos de incidentes em grandes eventos	SVM	0,99
Ferreira (2018)	Peças jurídicas	Redes Neurais	0,94
Costa (2019)	<i>Fake News</i>	Regressão Logística	0,89
Batista (2017)	Itens de Notas fiscais	Naive Bayes	0,88
Oliveira <i>et al.</i> (2018)	Enunciado de questões	Redes Neurais	0,87

Ferreira (2012)	Atividades econômicas de empresas	de	Naive Bayes	0,81	
Marumo (2018)	<i>Fake News</i>		Redes Neurais	0,79	
Rolim <i>et al.</i> (2017)	Postagens educacionais	em	fóruns	Redes Neurais	0,76

Fonte: Elaboração própria.

Observa-se uma grande variação da acurácia obtida pelos trabalhos, com alguns beirando à perfeição (99% de acertos) e outros chegando a 76% de acerto.

Fato é que aqueles que obtiveram melhores resultados foram aqueles realizados em contextos onde algumas palavras indicam fortemente qual a classificação do texto. O trabalho de Lima & Pinto (2016), que obteve 99% de acurácia, por exemplo, trata de classificação de um conjunto de texto muito específicos que descrevem protestos, acidentes de trânsito, entre outros tipos de manifestações.

Outro trabalho que obteve acurácia excelente foi o de Ferreira (2018) que classificou peças jurídicas. Sobre peças jurídicas, destaca-se que o uso de um sub vocabulário próprio, conhecido popularmente como “juridiquês”, diminui a quantidade de palavras que serão utilizadas nessas peças. Além disso, existe forte relação entre algumas palavras e algumas peças (por exemplo, toda sentença vai ter a palavra “decido”).

Dessa forma, em conjuntos de textos que seguem alguma estruturação, se utilizando de vocabulários específicos e existe uma correlação forte entre algumas palavras e uma determinada classificação, verifica-se que os modelos de AI terão mais facilidade para classificar corretamente os dados de entrada.

Observa-se que os trabalhos de classificação de texto em português que obtiveram acurácia menor do que 90% apresentam esta característica: os textos de entrada versam sobre assuntos diversos (*fake news* podem ser sobre qualquer área do governo, postagem em fóruns educacionais podem ser sobre qualquer área do conhecimento etc.).

Uma possível causa para que apenas 25% dos trabalhos de classificação em língua portuguesa tenham alcançado mais de 90% de acurácia pode ser a própria língua portuguesa. Algumas estratégias de pré-processamento disponíveis em inglês, como a stemização e a lematização, ainda estão em desenvolvimento para fornecer um melhor *corpus* da língua portuguesa.

Outra avaliação que pode ser feita aqui é sobre o algoritmo de ML utilizado pelos trabalhos de classificação de textos em português. Apesar de serem testados diversos algoritmos nesses trabalhos, 50% obtiveram melhores resultados com a utilização de redes neurais.

Provavelmente, o classificador não linear baseado em pesos e funções de ativação criado pelas redes neurais esteja melhor adaptado à diversidade de regras gramaticais e sinônimos da língua portuguesa.

3.8 RESULTADOS

3.8.1 Validação Cruzada com os Dados de Treinamento

Para escolha da melhor configuração de hiperparâmetros de cada modelo, utilizou-se o conjunto de treinamento das recomendações para treinamento e validação de cada configuração.

Para validação de cada configuração, utilizou-se a validação cruzada *10-fold*. Essa técnica consiste em dividir o conjunto de treinamento em dez partes aleatórias, sendo utilizadas nove dessas partes como conjunto de treino da instância de modelo e uma parte como conjunto de testes. Cada instância de modelo foi treinada dez vezes, onde cada execução se utilizou uma das dez partes como conjunto de testes.

Os resultados dessa etapa estão apresentadas na Tabela 2.

Tabela 2 - Resultados obtidos com a variação de parâmetros dos modelos treinados, obtidos por meio de validação cruzada 10-k fold.

Modelo	Configuração de Hyperparâmetros	Acurácia da Configuração*
Regressão Logística	solver=lbfgs	0,68 ± 0,01
Regressão Logística	solver=newton-cg	0,68 ± 0,01
Regressão Logística	solver=liblinear	0,68 ± 0,01
Regressão Logística	solver=sag	0,68 ± 0,01
Regressão Logística	solver=saga	0,68 ± 0,01
Naive Bayes	alpha=1.0	0,66 ± 0,01
Naive Bayes	alpha=0	0,67 ± 0,01
Support Vector Machines	kernel=linear	0,69 ± 0,01

Support Vector Machines	kernel=rbf	0,57 ± 0,01
Support Vector Machines	kernel=poly	0,53 ± 0,01
Support Vector Machines	kernel=sigmoid	0,57 ± 0,01
Rede Neural MLP	solver=lfbgs, alpha = 1e-5, hidden_layer_sizes = (5, 2)	0,70 ± 0,01
Rede Neural MLP	solver=lfbgs, alpha=1e-5	0,70 ± 0,02
Rede Neural MLP	solver=lfbgs, alpha=1e-5, activation=logistic	0,52 ± 0,12
Rede Neural MLP	solver=lfbgs, alpha=1e-5, activation=tanh	0,71 ± 0,02
Rede Neural MLP	alpha=1e-5, activation=tanh	0,76 ± 0,02
Rede Neural MLP	solver=sgd, alpha=1e-5, activation=tanh	0,66 ± 0,01
Rede Neural MLP	alpha=1e-5, activation='tanh', learning_rate_init = 0.01	0,86 ± 0,01
Rede Neural MLP	alpha=1e-5, activation='tanh', learning_rate_init = 0.1	0,85 ± 0,03

*Acurácia média das dez execuções da instância do modelo e desvio padrão multiplicado por dois. Fonte: Elaboração própria.

3.8.2 Regressão Logística

A variação dos hiperparâmetros da RL não resultou em variação dos resultados obtidos na validação cruzada. Assim, executou-se o conjunto de teste na instância de RL cujo treinamento foi o mais rápido (solver=lfbgs), o que gerou a matriz de confusão apresentada no Quadro 3.

Quadro 3: Matriz de Confusão do melhor modelo de Regressão Logística.

	Não atendida	Atendida
Não Atendida	2.272	1.864
Atendida	2.866	6.193

Fonte: Elaboração Própria.

Com relação às métricas de avaliação da RL, a Figura 8 apresenta os valores obtidos.

Figura 8: Precisão, *recall* e *f1-score* obtidos na etapa de testes do melhor modelo de Regressão Logística.

	precision	recall	f1-score	support
0	0.44	0.55	0.49	4136
1	0.77	0.68	0.72	9059
micro avg	0.64	0.64	0.64	13195
macro avg	0.61	0.62	0.61	13195
weighted avg	0.67	0.64	0.65	13195

Fonte: Elaboração Própria.

3.8.3 Multinomial Naive Bayes

O melhor modelo de NB gerado a partir do treinamento com validação cruzada e aplicado sobre o conjunto de testes, gerou a matriz de confusão apresentada no Quadro 4.

Quadro 4: Matriz de Confusão do melhor modelo de Multinomial Naive Bayes.

	Não atendida	Atendida
Não Atendida	2.142	1.994
Atendida	2.513	6.546

Fonte: Elaboração Própria.

Com relação às métricas de avaliação do referido modelo, a Figura 9 apresenta os valores obtidos.

Figura 9 - Precisão, *recall* e *f1-score* obtidos na etapa de testes do melhor modelo de Multinomial Naive Bayes.

	precision	recall	f1-score	support
0	0.46	0.52	0.49	4136
1	0.77	0.72	0.74	9059
micro avg	0.66	0.66	0.66	13195
macro avg	0.61	0.62	0.62	13195
weighted avg	0.67	0.66	0.66	13195

Fonte: Elaboração Própria.

3.8.4 Support Vector Machine

O melhor modelo de SVM gerado a partir do treinamento com validação cruzada e aplicado sobre o conjunto de testes, gerou a matriz de confusão apresentada no Quadro 5.

Quadro 5: Matriz de Confusão do melhor modelo de *Support Vector Machine*.

	Não atendida	Atendida
Não Atendida	2.256	1.880
Atendida	2.896	6.163

Fonte: Elaboração Própria.

Com relação às métricas de avaliação do referido modelo, a Figura 10 apresenta os valores obtidos.

Figura 10 - Precisão, *recall* e *f1-score* obtidos na etapa de testes do melhor modelo de *Support Vector Machine*.

	precision	recall	f1-score	support
0	0.44	0.55	0.49	4136
1	0.77	0.68	0.72	9059
micro avg	0.64	0.64	0.64	13195
macro avg	0.60	0.61	0.60	13195
weighted avg	0.66	0.64	0.65	13195

Fonte: Elaboração Própria.

3.8.5 Redes Neurais

O melhor modelo de RN MLP gerado a partir do treinamento com validação cruzada e aplicado sobre o conjunto de testes, gerou a matriz de confusão apresentada no Quadro 6.

Quadro 6: Matriz de Confusão do melhor modelo de Rede Neural MLP.

	Não atendida	Atendida
Não Atendida	1.494	2.642
Atendida	1.464	7.595

Fonte: Elaboração Própria.

Com relação às métricas de avaliação do referido modelo, a Figura 10 apresenta os valores obtidos.

Figura 11 - Precisão, *recall* e *f1-score* obtidos na etapa de testes do melhor modelo de Rede Neural.

	precision	recall	f1-score	support
0	0.51	0.36	0.42	4136
1	0.74	0.84	0.79	9059
micro avg	0.69	0.69	0.69	13195
macro avg	0.62	0.60	0.60	13195
weighted avg	0.67	0.69	0.67	13195

Fonte: Elaboração Própria.

3.8.6 Comparação entre os Modelos

A Tabela 3 apresenta uma comparação entre os modelos elaborados no presente trabalho:

Tabela 3: Comparação da precisão e *recall* médios (*micro avg*) dos modelos desenvolvidos no presente trabalho.

	Regressão Logística	Naive Bayes	Support Vector Machine	Rede Neural MLP
Precisão	e 0,64	0,66	0,64	0,69
Recall				

Fonte: Elaboração própria.

Assim, fica evidente que o melhor modelo desenvolvido foi o que se utiliza da rede neural MLP.

4 CONCLUSÃO

É possível construir um preditor das recomendações emitidas pela CGU, tomando como entrada apenas o texto da recomendação. Essa tarefa foi realizada no presente trabalho, sendo elaborada, parametrizada e treinada uma rede neural que, durante a etapa de testes, obteve uma precisão de 69% no que concerne a se uma determinada recomendação seria atendida ou cancelada / não monitorada.

Esse percentual de acerto não foi maior possivelmente pela dificuldade da base de recomendações, as quais tratam de diversos assuntos e estão escritas em uma linguagem de difícil aprendizado, o português.

Seguindo o modelo do CRISP-DM, a próxima etapa é a implantação desse preditor em modo de produção. Uma das estratégias que será avaliada é a produção de um módulo de avaliação das recomendações escritas pelos auditores *on-time*, no e-Aud, sistema de gestão das auditorias efetuadas pela CGU.

Além da implantação do modelo, uma nova iteração no CRISP-DM deve ser planejada com o objetivo de melhorar a precisão e *recall* do modelo. Entre as estratégias que devem ser pensadas, registra-se:

- Incluir no a stemização e a lematização no pré-processamento dos textos da recomendação;
- Testar outros modelos de redes neurais, como as convolucionais e as recorrentes; e
- Utilizar hardware mais potente no processo de treino e testes, realizando um *grid search* mais extensivo dos hiperparâmetros da rede neural.

5 REFERÊNCIAS

ARANTES, Rodrigo Eloy. **O controle interno como apoio à gestão dos órgãos públicos? Uma análise sob a perspectiva dos gestores federais.** 2016. Dissertação (Mestrado em Políticas Públicas) – Instituto de Pesquisa Econômica Aplicada, Distrito Federal, 2016.

ARAÚJO, Fábio da Silva. **Controle Interno no Poder Executivo Federal: Um estudo exploratório quanto à percepção dos auditores internos das Instituições Federais de Ensino (IFE's) sobre a atuação da Controladoria-Geral da União (CGU).** 2007. Dissertação (Mestrado em Ciências Contábeis) – UnB, UFPB, UFPE e UFRN, Recife, 2007.

AVRITZER, Leonardo. **Impasses da Democracia no Brasil.** Rio de Janeiro: Ed. Civilização Brasileira, 2016.

AZEVEDO, Alexandre Pandino. **Efeitos da Auditoria na melhoria da Gestão Pública: a implementação das recomendações de auditoria, pode ser uma ferramenta estratégica na melhoria da gestão pública?** 2016. Trabalho de Conclusão de Curso (Especialização) – Escola Nacional de Administração Pública, Brasília, 2016.

BATISTA, Rodrigo de Abreu. **Classificação automática de códigos NCM utilizando o algoritmo Naive Bayes.** 2017. Trabalho de Conclusão de Curso (especialização) – Universidade de Santa Cruz do Sul, Porto Alegre, 2017.

BOSE, Bernhard E., GUYON, Isabelle M. & VAPNIK, Vladimir N. A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop of Computational Learning Theory. Pittsburgh: ACM Press, 1992, p. 144-152.

BRINK, Henrik.; RICHARDS, Joseph. W. & FETHEROLF, Mark. **Real-World Machine Learning.** Shelter Islands: Manning, 2016.

CONTROLADORIA-GERAL DA UNIÃO. **Manual de Integridade Pública e Fortalecimento da Gestão: Orientações para o Gestor Municipal** 3. ed. Brasília: CGU, 2013.

CONTROLADORIA-GERAL DA UNIÃO. **Sistema Monitor – Manual de Monitoramento das Ações de Controle – Versão do Gestor**. Versão 6. Brasília: CGU, 2018.

CHAPMAN, Pete *et al.* **CRISP-DM 1.0 step-by-step data mining guide**. 2000. Disponível em <https://www.the-modeling-agency.com/crisp-dm.pdf>. Acessado em 02/03/2020.

COMMITTEE OF SPONSORING ORGANIZATIONS OF THE TREADWAY COMMISSION. **Controle Interno – Estrutura Integrada: sumário executivo**. Jersey City: AICPA, 2013.

COSTA, Lucas Gomes. **Classificação de fake news utilizando algoritmos de aprendizado de máquina e aprendizado profundo**. 2019. Monografia (Graduação) – Universidade Federal de Roraima, Boa Vista, 2019.

DAVILA, Marcos; OLIVEIRA, Marcelo. **Conceitos e Técnicas de Controle Interno de Organizações**. São Paulo: Nobel, 2002.

DIAS, Valmir Gomes. **Auditorias de obras públicas: variáveis determinantes do atendimento de recomendações da Controladoria Geral da União**. 2018. Dissertação (Mestrado) – Universidade Federal da Bahia, Salvador, 2018.

FERREIRA, Maria da Conceição da Silva. **Classificação Hierárquica da Atividade Económica das Empresas a partir de Texto da Web**. 2012. Tese (Mestrado em Análise de Dados) – Universidade do Porto, Porto, 2012.

FERREIRA, Marcelo Herton Pereira. **Classificação de peças processuais jurídicas: Inteligência Artificial no Direito**. 2018. Trabalho de Conclusão de Curso (Graduação) – Universidade de Brasília, Gama, 2018.

GOLDBERG, Yoav. **Neural Network Methods for Natural Language Processing**. San Rafael: Morgan & Claypool, 2017.

HANASHIRO, Maíra. **Auditoria de TI na CGU: proposta de modelo de implementação de auditoria de tecnologia da informação no âmbito da Controladoria-Geral da União**. 2009. Relatório Técnico-Científico (Especialização) – Escola da AGU, Cefor, SFC/CGU e ISC/TCU, Brasília, 2009.

HEARST, Marti *et al.* Support vector machines. **Intelligent Systems and their Applications**, IEEE, v. 13, n. 4, p. 18–28. 1998.

HOTHO, Andrea *et al.* A brief survey of text mining. **LDV Forum - GLDV Journal for Computational Linguistics and Language Technology**, v. 20, n. 1, p. 19–62. 2005.

HAN, Jiawei; KAMBER, Micheline. **The Data Mining: Concepts and Techniques**. 2ª ed. São Francisco: Morgan Kaufmann Publishers, 2011.

KINGMA, Diederik P.; BA, Jimmy Lei. Adam: A method for stochastic optimization. **International Conference on Learning Representations**, v. abs/1412.6980. 2014

LIMA, André Accioly; PINTO, Renato Carlos. **Estudo Experimental de Aprendizado de Máquina para Desenvolvimento de um Classificador de Texto de Incidentes de Grandes Eventos**. 2016. Monografia (Graduação) – Universidade de Brasília, Brasília, 2016.

MADEIRA, André Aldo Dias. **A importância do Follow-up na Auditoria Interna em Portugal**. 2013. Dissertação (Mestrado em Auditoria) – Instituto Superior de Contabilidade e Administração de Lisboa, Lisboa, 2013.

MARUMO, Fabiano Shiiti. **Deep Learning para classificação de Fake News por sumarização de texto**. 2019. Trabalho de Conclusão de Curso (Graduação em Ciência da Computação) – Universidade Estadual de Londrina, Londrina, 2019.

MARX, César Augusto. **A CGU e a dualidade do papel do Controle Interno no Brasil**. 2015. Dissertação (Mestrado) – Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas, São Paulo, 2015

MITCHELL, Tom M. **Machine Learning**. McGraw-Hill, 1997.

MLODINOW, Leonard. **O Andar do Bêbado: Como o acaso determina nossas vidas**. Rio de Janeiro: Zahar, 2009.

MUNIZ, Frederico Perini. **A trajetória institucional do sistema de controle interno do poder executivo federal: a emergência e a evolução da Controladoria Geral da União**. 2017. Dissertação (Mestrado Profissional em Gestão Pública) – Universidade Federal do Espírito Santo, Vitória, 2017.

OLIVEIRA, Pedro Igor de Araújo *et al.* **Ferramenta de classificação de questões para auxílio ao aprendizado**. 2018. Projeto de Fim de Curso (Graduação) – Instituto Militar de Engenharia, Rio de Janeiro, 2018.

OLIVIERI, Cecília. **Política e burocracia no Brasil: o controle sobre a execução das políticas públicas**. 2008. Tese (Doutorado em Administração Pública e Governo) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2008.

PEDREGOSA, Fabian. *et al.* **Scikit-learn: Machine Learning in Python**. Journal of Machine Learning Research, v. 12, p. 2.825-2.830. 2011.

PETER, Maria da Glória Arrais; MACHADO, Marcus Vinícius Veras. **Manual de auditoria governamental**. 2 ed. São Paulo: Editora Atlas, 2014.

RASCHKA, S. **Single-Layer Neural Networks and Gradient Descent**. 2015. Disponível em: <https://sebastianraschka.com/Articles/2015_singlelayer_neurons.html>. Acesso em 02/03/2020.

REZENDE, Solange. O. et al. **O uso da mineração de textos para extração e organização não supervisionada de conhecimento**. Revista de Sistemas de Informação da FSMA, n. 7, pp. 7–21. 2011.

ROLIM, Vitor Belarmino *et al.* Utilização de técnicas de aprendizado de máquina para acompanhamento de fóruns educacionais. **Revista Brasileira de Informática na Educação**, 25(3), p. 112-130. 2017.

RUSSEL, Stuart; NORVIG, Peter. **Inteligência Artificial**. 3ª ed. Rio de Janeiro: Elsevier Editora, 2013

Salton, Gerard; Buckley, Christopher. Term-weighting approaches in automatic text retrieval. **Journal Information Processing and Management**, v. 24, cap. 5, pp. 513-523. 1988.

SANTOS, Igor Pedro Pinto. **Análise de Sentimento Usando Redes Neurais de Convolução**. 2017. Dissertação (Mestrado) – Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2007.

SAYAD, Saed. **An Introduction to Data Mining**. Disponível em: <<http://www.saedsayad.com/>>. Acesso em 02/03/2020.

SILVA, Fabrício Ricardo Do Rêgo Alves. **Recomendações do Controle Interno no Poder Judiciário no período de 2012 a 2014**. 2016. Trabalho de Conclusão de Curso (Especialização) Centro Universitário IESB, Brasília, 2016.