

**Desenvolvimento de um Sistema de Recomendação de Candidatos à Cargos Públicos  
Utilizando Técnicas de Aprendizagem de Máquina e Filtragem de Informação**

**Orientador(a): Fábio Santos da Silva**  
**Co-orientador(a):**  
**Bolsista/Voluntário(a): Lucas Pereira Reis**

## 1. Introdução

Eleição é todo processo onde um grupo de pessoas escolhe seus representantes de algum cargo público através de votações. As eleições no Brasil são realizadas através do voto direto, secreto e obrigatório [1].

As eleições ocorrem a cada quatro anos, nos anos pares. À exceção do cargo de senador, onde seu mandato possui duração de oito anos, os demais cargos seguem o padrão de quatro anos. Os cargos eletivos são disputados em dois grupos, da seguinte forma:

- Eleições federais e estaduais – para os cargos de: Presidente da República (e vice), Senador, Deputado Federal, Governador (e vice) e Deputado Estadual.
- Eleições municipais – para os cargos de Prefeito (e vice) e Vereadores.

Os candidatos a qualquer cargo são filiados a algum dos mais de 30 partidos políticos legalizados existentes no país, cada um com uma ideologia política. Todos os partidos recebem recursos do fundo partidário, acesso aos meios de comunicação (rádio e TV), e direito ao horário eleitoral durante as campanhas [1].

Segundo a Constituição Federal, em seu artigo 14, o voto é facultativo para os analfabetos, aos maiores de 70 anos e para os maiores de 16 anos e menores de 18 anos. É obrigatório para os cidadãos entre 18 e 70 anos, sendo necessário justificar a ausência em qualquer seção eleitoral, no dia da eleição, sob pena de multa [1]. Com base no que foi apresentado, percebe-se a importância do voto e a importância das pessoas escolherem seus representantes através dos seus próprios interesses e opiniões.

No Brasil a cada eleição muitos políticos são candidatos para inúmeros cargos, gerando uma dificuldade dos eleitores em analisar todas as propostas de todos os candidatos para votar no mais apropriado. Esta dificuldade de escolha de uma opção diante de uma diversidade é conhecida na literatura como o problema da Sobrecarga de informações [3]. Uma das consequências deste problema é que o eleitor acaba escolhendo o candidato por meio de amizade, interesse, recomendações de amigos, por *marketing* através de TVs e rádios [2].

Nas eleições presidenciais de 2014 mais de 27,44% dos votos válidos não foram direcionados para candidatos, mostrando que muitos eleitores não se sentiam representados por nenhum deles. [5]

### 1.1 Justificativa

Diante do cenário apresentado, se justifica o desenvolvimento de um sistema de auxílio ao eleitor durante as eleições, para direcionar os eleitores a realizarem uma escolha apropriada. É possível que talvez em 2022, mais da metade dos eleitores se absterão de votar, anularão o voto ou votarão nulo. [5]

Atualmente, existem algumas iniciativas para ajudar na escolha, aplicativos como Voto x Veto, Repolítica, Ficha Suja, Proteste [6], o Google também desenvolveu um sistema para auxílio [7]. Porém, analisando cada um separadamente, percebe-se que nenhum destes sistemas exploram o conceito de recomendação personalizada de candidatos, apenas exibem as propostas, as fichas, dados estatísticos, gráficos sobre o desempenho do candidato, e informações diversas do candidato.

Nesta proposta do projeto de forma diferente explora-se a utilização de Sistemas de Recomendação, o qual é utilizado para gerar recomendações válidas de itens que possam interessar aos usuários [3]. Para este projeto, o termo item refere-se a uma pessoa, o candidato ao cargo público.

Sistemas de Recomendação utilizam técnicas de Aprendizagem de Máquina (AM) e Recuperação de Informação (RI) para gerar recomendações baseadas em algum tipo de filtragem [3][4]. O sistema de recomendação é um tema que tem despertado grande atenção da comunidade acadêmica que se reúne em importantes eventos relacionado ao tema como: SIGIR [8], RECSYS [9] e CIKM [10].

Neste projeto, pretende-se estudar e empregar técnicas de Aprendizagem de Máquina (classificação e clusterização) e Filtragem de Informação (FI) para o desenvolvimento de um Sistema de Recomendação de Candidatos a Cargos Públicos

Desta forma, para o problema da Sobrecarga de informações no âmbito das eleições, este projeto visa, justamente, desenvolver um sistema de recomendação de candidatos à órgãos públicos para auxiliar os eleitores durante a escolha.

A recomendação do sistema será baseada em um processo de filtragem que usa Aprendizagem de Máquina e Filtragem de Informações para mensurar o grau de relevância do candidato com relação aos temas: educação, saúde, segurança, economia, entre outras correlacionando com os temas que o eleitor prioriza por meio de uma escala, por exemplo, de 1 a 5 que o mesmo indica ao sistema.

Deste modo, o sistema deverá gerar uma recomendação de candidatos apresentando um ranking que coloca no topo os candidatos que possuem maior valor de similaridade com as prioridades do eleitor.

## **2. Objetivo(s)**

Os objetivos estabelecidos para este trabalho são os seguintes:

### **2.1 Objetivo Geral**

Esta proposta de projeto possui como objetivo geral desenvolver um sistema de recomendação de candidatos à cargos públicos utilizando técnicas de Aprendizagem de Máquina e Filtragem de Informação. O sistema será chamado de EmQuemVotar.

### **2.2 Objetivos Específicos**

1. Realizar revisão bibliográfica de conteúdos relacionados às áreas de Sistemas de Recomendação, técnicas de Filtragem de Informação e Aprendizagem de Máquina;
2. Investigar e pesquisar tecnologias apropriadas para o desenvolvimento do projeto;
3. Realizar a aquisição de dados sobre políticos através da internet.
4. Realizar pré-processamento dos dados, gerando o *dataset*;

5. Projetar a arquitetura do sistema, seus componentes que implementam as técnicas de aprendizagem de máquina e filtragem de informação;
6. Desenvolver um protótipo para *web* integrando-o com o sistema gerado.
7. Realizar testes para avaliar a qualidade das recomendações geradas pelo sistema

### 3. Metodologia

A metodologia empregada para alcançarmos todos os objetivos específicos será feita da seguinte forma:

Uma revisão bibliográfica ampla sobre Sistemas de Recomendação, técnicas de Filtragem de Informação e Aprendizagem de Máquina, tanto a fundamentação teórica quanto a implementação de seus algoritmos, será realizada. Esta revisão contempla o objetivo específico (1).

No objetivo específico (2), será feito pesquisas sobre quais serão as ferramentas a serem utilizadas para os próximos objetivos específicos, por exemplo: como será implementado a aquisição de dados, qual técnica de Aprendizagem de Máquina a ser considerada e utilizada, que tipo de técnica de filtragem utilizar, como implementar um sistema de recomendação e qual *framework* para Aprendizagem de Máquina e Filtragem de Informação utilizar.

A partir do objetivo específico (3) inicia-se de fato, a implementação do projeto. Inicialmente, através da internet, será feita a aquisição de dados de cada político a ser considerado no projeto utilizando-se da ferramenta pesquisada no objetivo específico (2), retirando apenas as informações necessários para o modelo futuramente, por exemplo o interesse do político nas áreas de tecnologia, saúde, segurança etc.

No objetivo específico (4), após realizar toda a coleta de dados, será feito um pré-processamento dos dados (organizá-los, definir atributo-alvo, excluir ruídos e dados faltosos, etc.) para que no final o *dataset* do projeto esteja pronto.

No objetivo específico (5), o *dataset* obtido anteriormente será utilizado para o treinamento do modelo baseado na técnica de Aprendizagem de Máquina definida no objetivo específico (2), e em seguida, para os testes do mesmo. Após verificar que o modelo se encontra em um estado aceitável para o projeto, será implementado a Filtragem de Informação, deixando o sistema pronto para a próximo etapa.

No objetivo específico (6), será desenvolvido e implementado um protótipo do sistema para o *web* utilizando a tecnologia estudada no objetivo específico (2). Finalmente, testes serão realizados para avaliar a qualidade das recomendações geradas pelo sistema. Para isso, serão empregadas métricas para avaliar este tipo de sistema.

### 3 Resultados esperados

As principais contribuições esperadas ao final do desenvolvimento dessa pesquisa são as seguintes:

1. Auxiliar o eleitor a abstrair a complexidade da interpretação de dados estatísticos, gráficos, normalmente usados em portais sobre políticos por meio de recomendação personalizada;
2. Diminuir a dificuldade de escolha de candidatos à cargos públicos;

- 3 A recomendação apresentar candidatos satisfatórios para o usuário.
4. O sistema ser utilizado por uma grande parcela de eleitores;
5. Publicação de artigos em periódicos e eventos acadêmicos relacionados ao tema.

#### 4 Cronograma de execução

	2018					2019						
Atividades	08	09	10	11	12	01	02	03	04	05	06	07
1. Estudo sobre sistemas de recomendação e trabalhos relacionados	X											
2. Estudo de técnicas de Aprendizagem de Máquina e Filtragem de Informações		X	X									
3. Identificação dos principais requisitos do sistema, projeto da arquitetura e componentes				X								
4. Desenvolvimento de componentes de filtragem de informação e aprendizagem de máquina					X	X	X					
5. Desenvolvimento do Frontend do Sistema e Backend								X	X			
6. Testes e Avaliação do sistema										X	X	
7. Elaboração de Relatórios e Artigos	X	X	X	X	X	X	X	X	X	X	X	X

#### 5 Referências

- [1] PACIEVITCH, Thais. *Eleições no Brasil*, Infoescola. Disponível em: <<https://www.infoescola.com/direito/eleicoes-no-brasil/>>. Acessado em: 22 de abril de 2018.
- [2] BURGARELLI, Rodrigo. *Voto inconsciente: o que o brasileiro leva em conta para decidir seu candidato a prefeito?*, Estadão. Disponível em: <<http://alias.estadao.com.br/noticias/geral,voto-inconsciente-o-que-o-brasileiro-leva-em-conta-para-decidir-seu-candidato-para-prefeito,10000077956>>. Acessado em: 22 de abril de 2018.

- [3] ADOMAVICIUS, G.; TUZHILIN, A. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. In: IEEE Transactions on Knowledge and Data Engineering, v.17, n.6, p.734-749, 2005
- [4] MCCREA, N. (2015). An Introduction to Machine Learning Theory and Its Applications: A Visual Tutorial with Examples. Acesso em Dezembro de 2017, disponível em l: <https://www.toptal.com/machine-learning/machine-learning-theoryan-introductory-primer>
- [5] FAZZIO, Luiz. *Os brasileiros querem ter em quem votar*, Estadão. Disponível em: <<http://politica.estadao.com.br/blogs/fausto-macedo/os-brasileiros-querem-ter-em-quem-votar/>>. Acessado em: 26 de abril de 2018.
- [6] ÉPOCA, *10 Aplicativos e sites que ajudam a decidir em quem votar*, Globo. Disponível em : <<https://epocanegocios.globo.com/Informacao/Acao/noticia/2014/10/10-aplicativos-e-sites-que-ajudam-decidir-em-quem-votar.html>>. Acessado em: 27 de abril de 2018.
- [7] PENATTI, Giovana. *Ainda não sabe em quem votar? O Google ajuda*, Tecnoblog. Disponível em: <<https://tecnoblog.net/166224/ainda-nao-sabe-em-quem-votar-google-ajuda/>>. Acessado em: 27 de abril de 2018.
- [8] SIGIR. Disponível em <<http://sigir.org/sigir2018/>>. Acesso em: 2018-04-23
- [9] RECSYS. Disponível em <<https://recsys.acm.org/recsys18/>>. Acesso em: 2018-04-23
- [10] CIKM. Disponível em < <http://cikm2018.fudan.edu.cn/>>. Acesso em: 2018-04-23