CONCLUSÃO

Na busca de algumas instituições de abusarem da psicologia humana para divulgar informações falsas, recentemente na eleição presidencial estadunidense, que deixou o tema em voga, mas também vem de muitos anos antes, sendo a mídia uma massa de manipulação, surge o auxílio da Ciência de Dados de maneira a utilizar do reconhecimento de padrões para detectar Fake News e parar a sua divulgação.

Nesta monografia foi discutido o panorama histórico entre os dois temas, como a as notícias falsas sempre foram um problema para a humanidade e usadas como uma massa de manobra, e o surgimento de tecnologias que tornaram possível a ampla divulgação da informação, tornando mais difícil o trabalho de quem queria criar mentiras, mas tornando mais rápido o processo de sua divulgação. Ao mesmo tempo, mostramos o surgimento da Ciência de dados e como os aspectos que viriam a ser responsáveis pela área de reconhecimento de padrões foram se desenvolvendo ao longo dos anos até abordarem o problema diretamente desde 2016.

Foi discutido uma definição mais formal do que é Fake News e como ela é debatida no meio acadêmico e pode gerar diversos problemas ao se desenvolver pesquisas e algoritmos no meio, e também a responsabilidade dos cientistas ao fazerem tais pesquisas por se tratar de um tema com diversas nuances que podem categorizar a detecção em censura.

Foram abordadas diferentes estratégias usadas hoje no meio para a detecção efetivas das Fake News, que estão divididas em três grandes campos: A análise do conteúdo, das fontes e a mudança nos incentivos que pessoas ou sites tem para compartilhar. Na primeira, onde analisa-se o conteúdo da notícia ou postagem em si, existem diferentes formas sendo feitas e testadas para fazer isso, entre elas as flags, os algoritmos óbvios e os baseados nas Wikipedia Hoaxes, que é um sistema de heurística estabelecido pela Wikipedia para determinar artigos falsos em sua database. Na detecção de fontes vimos que o maior foco no meio é a detecção de sistemas de bots, pois já existem uma boa base de testes e de pesquisas na área, e o desafio é usar desses sistemas para fazer um paralelo com a detecção especificamente de Fake News. Vimos então duas das maiores contribuições nesse meio, o BotOrNot e a Lei de Benford. Por fim, a modificação de incentivos que busca alterar a forma gratificante que Fake News são compartilhadas, tanto para sites que lucram ou querem fazer algum tipo de propaganda em cima dela, tanto para pessoas que acreditam naquela informação, e como alterar esse meio atual de forma a não ter esse incentivo.

Além de muitas estratégias dentre esses três pontos abordadas no capítulo 3, também vimos como um amplo maior as maiores dificuldades que a Ciência de Dados encontra nesse contexto, sendo a falta de dados numerosos e definições incertas de escopo do problema. Desafios esses que têm tendência de serem mais facilmente resolvidos com o tempo, por ser uma abordagem muito nova a ideia de detectar Fake News com reconhecimento de padrões. Foi mostrado como parcerias com redes sociais ou sistemas que já abordavam problemas semelhantes parece ser a tendência pela quantidade de dados disponíveis e pesquisas já feitas na área que podem ser usadas como um apoio.