**Sentiment Analysis in Czech Social Media Using Supervised Machine**

**Learning**

**Contextualização**

Em que área do conhecimento o artigo se encaixa? Qual o contexto geral da utilização/aplicação dessa pesquisa? Qual é o problema?

Pesquisa com a finalidade em estabelecer um grande campo em métodos para análise de sentimento em mídias sociais tchecas, já quem em inglês, chinês ou espanhol esse campo já tem uma longa história.

**Trabalhos relacionados**

a. Soluções existentes: quais os artigos que já investigaram esse problema?

Temos trabalhos com o da Kateřina Veselovská “Sentence-level sentiment analysis in czech” que apresentou um trabalho inicial de análise de sentimento tcheco e do Josef Steinberger que fez uma triangulação aproximada de sentimentos tchecos em “Creating sentiment dictionaries via triangulation”.

b. Solução mais próxima: qual é a melhor e/ou solução mais próxima ao problema apresentado?

Logo no abstrato do texto vemos que tem muitas pesquisas em inglês, chinês e espanhol, porém em tcheco não ocorreu a condução de nenhuma pesquisa sistemática.

**Gap**

Qual a principal/is limitações da(s) solução(ões) existente(s)? O que não foi investigado? Quais aspectos do problema ainda não estão claros? Por que as abordagens existentes podem ser falhas?

Necessidade de uma linguagem natural de processamento que precisam lidar com alta flexão e uma rica sintaxe na linguagem tcheca e uma necessidade de um conjunto de dados bem documentados e de livre acesso.

**Hipótese ou questão de pesquisa**

Dado a limitação / problema listado anteriormente, o autor argumenta que há uma forma de fazer diferente que talvez resolva o problema, ou ainda atinja resultado melhor ou que investigue algum aspecto?

Para conseguir analisar sentimentos existem duas abordagens que são “dictionary-based” e “machine learning-based” .

**Objetivo**

O autor do artigo irá utilizar modelos de machine learning como “N-gram features”, “N-gram features”, “POS-related features”, ”Emoticons” e “Delta TFIDF variants for binary scenarios”, para se o sentimento é positivo, negativo ou neutro.

**Metodologia:**

a. Qual o tipo de revisão realizada: narrativa, sistemática, integrativa, ou outro tipo?

Foi utilizado a revisão sistemática.

b. Quais são as amostras, as variáveis, tipo de pesquisa e as métricas utilizadas?

As amostras foram 10.000 posts aleatórios em tcheco de 9 diferentes páginas no Facebook onde essas bases tinha muitos fãs tchecos.

c. Como foi realizada a coleta e análise de dados?

Foi selecionada as páginas no Facebook onde eram top tchecos de acordo com a estatística do SocialBakers.

**Viabilidade De Detecção De Câncer De Mama Através De Rede Neural Convolucional Em Mamografias**

**Contextualização**

Em que área do conhecimento o artigo se encaixa? Qual o contexto geral da utilização/aplicação dessa pesquisa? Qual é o problema?

O estudo tem por finalidade conseguir fazer o diagnóstico mais precoce de câncer de mama, pois quando antes temos o diagnóstico maior é o índice de cura da doença.

**Trabalhos relacionados**

a. Soluções existentes: quais os artigos que já investigaram esse problema?

No MIT (Massachusetts Institute of Technology) foi obtido um nível de 90% de assertividade com técnicas de deep learning.

b. Solução mais próxima: qual é a melhor e/ou solução mais próxima ao problema apresentado?

Foram utilizados algoritmos de redes neurais artificiais e redes neurais convolucionais.

**Gap**

Qual a principal/is limitações da(s) solução(ões) existente(s)? O que não foi investigado? Quais aspectos do problema ainda não estão claros? Por que as abordagens existentes podem ser falhas?

Pelo baixo conjunto de imagens utilizando 5.500 de amostragem se pode obter uma baixa assertividade.

**Hipótese ou questão de pesquisa**

Dado a limitação / problema listado anteriormente, o autor argumenta que há uma forma de fazer diferente que talvez resolva o problema, ou ainda atinja resultado melhor ou que investigue algum aspecto?

A utilização de redes neurais artificiais e redes neurais convolucionais.

**Objetivo**

Conseguir através de redes neurais um acurácia significativa na identificação de câncer de mama.

**Metodologia:**

a. Qual o tipo de revisão realizada: narrativa, sistemática, integrativa, ou outro tipo?

Foi utilizado a revisão sistemática.

b. Quais são as amostras, as variáveis, tipo de pesquisa e as métricas utilizadas?

As amostras foram mais de 55.000 imagens, onde se tinha diversas métricas.

c. Como foi realizada a coleta e análise de dados?

Foi utilizada imagens de mamografias de banco de dados abertos como o Digital Database for Screening Mammography (DDSM - Mammography) da empresa Cancer Image Archives