

**RECONHECIMENTO DE DADOS
SENSÍVEIS NOS RELATÓRIOS DA
OUVIDORIA GERAL DO ESPÍRITO
SANTO**



CONCEITOS PRÉVIOS

- **Dados sensíveis** no contexto deste trabalho são todas informações pessoais dos cidadãos, conforme a legislação vigente e os padrões de trabalho da **Ouvidoria Geral do Estado**.
- A Ouvidoria Geral do Estado é um setor da **Secretaria de Controle e Transparência**, órgão central do Sistema de Controle Interno do Espírito Santo.



PROBLEMAS

Dados sensíveis informados inadvertidamente pelos cidadãos.

Ao preencherem o campo livre do sistema de Ouvidoria (e-Ouv) os cidadãos informam dados pessoais em suas solicitações.

Necessidade de transparência.

A Ouvidoria Geral precisa periodicamente produzir relatórios com o trâmite dos atendimentos realizados por ela e pelas Ouvidorias setoriais, sem, contudo, exibir os dados sensíveis dos cidadãos.

Lentidão na divulgação e análise dos relatórios.

Atualmente a verificação e exclusão dos dados sensíveis é feita manualmente, fato que consome um tempo considerável dos servidores envolvidos.



PROBLEMAS

Exemplo de uma manifestação feita pelos cidadãos (dados fictícios):

Meu nome é João Pereira da Silva, sou servidor público da SEGER com número funcional 2765478, meu CPF é 030.258.741-98, moro na rua Itacibá em Itaparica, Vila Velha. Estou com problema em receber a pensão da minha esposa, o nome dela é Maria Pereira da Silva e nosso telefone é (27) 981237415.



PROBLEMAS





SOLUÇÃO PROPOSTA

Utilizar uma abordagem híbrida baseada em expressões regulares e Reconhecimento de Entidades Nomeadas.

A solução deve receber o Relatório de Ouvidoria, reconhecer os dados sensíveis e gerar um novo relatório sem os dados sensíveis.

Por que usar o Reconhecimento de Entidades Nomeadas?

Reconhecer dados sensíveis na forma de Entidades Nomeadas se mostrou a abordagem mais efetiva, em função da variação lexical e baixa frequência das expressões textuais com dados sensíveis.

Aspectos tecnológicos.

O “aprendizado profundo” (*Deep Learning*) possibilitado pelas redes neurais tem apresentado excelentes resultados na detecção de Entidades Nomeadas. Neste trabalho foi utilizada a rede neural da biblioteca de PLN spaCy.



SOLUÇÃO PROPOSTA

Modelo Genérico X Modelo Específico

Optou-se pela construção de um corpus linguístico específico para o domínio do problema, pois nos testes iniciais esta abordagem se mostrou mais efetiva.

Um corpus linguístico é um conjunto de textos anotados com as informações necessárias ao algoritmo de *machine learning* da biblioteca spaCy, usada no presente trabalho.

Dificuldade de anotação.

Modelos para Reconhecimento de Entidades Nomeadas dependem diretamente da qualidade e representatividade dos dados de treinamento.

Todavia, o processo de anotação é manual e excessivamente trabalhoso, não sendo possível atingir um patamar adequado no prazo deste projeto.



SOLUÇÃO PROPOSTA

Exemplo de textos anotados para este trabalho (dados fictícios):

corpus_annotado = [

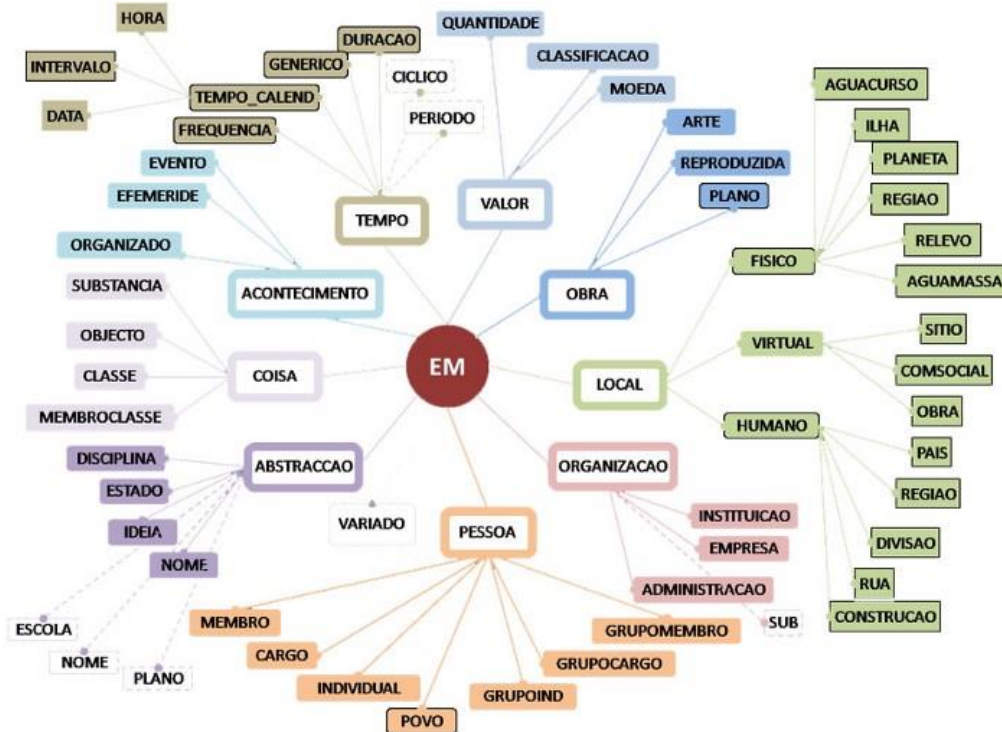
("Ola, bom dia, Meu nome e Luciana, sou aluna graduanda em licenciatura (biologia) na UFES. ",{"entities":[{"start":25,"end":32,"label":"CPER"}, {"start":84,"end":88,"label":"CORG"}]}),

("Sou Alaor Pereira, esposo da servidora Marcela Malory Alves, Pedagoga SEDU. ",{"entities":[{"start":3,"end":17,"label":"CPER"}, {"start":39,"end":59,"label":"CPER"}, {"start":71,"end":75,"label":"CORG"}]}))

]



SOLUÇÃO PROPOSTA



Exemplo de uma **Árvore de Entidades** criada a partir do corpus linguístico HAREM em português.



DADOS UTILIZADOS

Ano do Relatório de Manifestações	Manifestações de Ouvidoria	Textos Anotados para Treinamento	Textos Anotados para Teste e Validação
2018	3724	932	0
2019	3812	0	300
2020	3578	0	300



DADOS UTILIZADOS

No contexto deste trabalho as Entidades Nomeadas que devem ser reconhecidas são o que chamamos de “**Target**” ou “**Variáveis Dependentes**” no jargão da Ciência de Dados, enquanto os textos de entrada formam as chamadas “**Variáveis Independentes**” ou “**Features**”.



LIMITAÇÕES DOS DADOS

- A base de dados usada é homogênea, ou seja, as entidades nos textos não variam muito ao longo das Manifestações de Ouvidoria. Isto facilita o número de acertos nas classificações das entidades, mas impede afirmações sobre a generalização para outros domínios.
- Foram criadas entidades específicas para o domínio deste trabalho, tais como: CDOC (documentos diversos), CEND (endereço), CSAU (informações sobre saúde), CTEL (telefones), CMAIL (e-mail) e CPER (pessoas).
- O pré-processamento requerido limitou-se ao tratamento da estrutura dos blocos de texto, transformando cada Manifestação de Ouvidoria em uma sentença.



AVALIAÇÃO DO MODELO

MATRIZ DE CONFUSÃO	Palavras com Entidades Nomeadas	Palavras sem Entidades Nomeadas
Palavras previstas com Entidades Nomeadas	890	69
Palavras previstas sem Entidades Nomeadas	12	17.309

Matriz de confusão obtida com a base de validação contendo **18.280** palavras (*tokens*) e avaliando a presença ou não de entidades nomeadas.

A partir destes dados temos as seguintes medidas de desempenho para a detecção de entidades nomeadas:

- Acurácia: 99,56%
- Precisão: 92,81%
- Medida F: 95,65%



AVALIAÇÃO DO MODELO

TAXA DE ACERTO	Palavras classificadas corretamente	Palavras classificadas incorretamente
Número de palavras	831	71

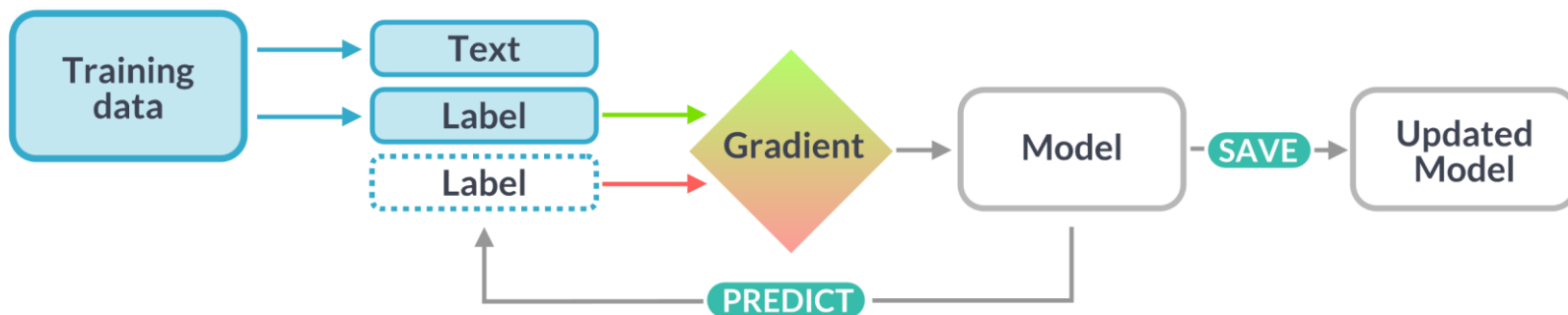
Taxa de acerto obtida com a base de validação contendo **902** palavras (*tokens*) previamente detectadas como Entidades Nomeadas. Neste cenário foi avaliado o índice de acerto na classificação .

Com os dados obtidos a taxa de acerto foi de **92,12%**.



INTERPRETAÇÃO DO MODELO

Considerando a utilização da biblioteca spaCy como recurso principal do presente trabalho, o modelo de treinamento e predição seguiu o padrão descrito pelos seus fabricantes.





COMPARAÇÃO COM OUTROS MODELOS

- Dada a especificidade deste trabalho e o uso da biblioteca spaCy não foi possível comparar o modelo elaborado com outros modelos durante a elaboração do projeto.
- Todavia, nada impede que outras soluções para Reconhecimento de Entidades Nomeadas sejam avaliadas no futuro e seus resultados comparados com os obtidos neste trabalho.



INSIGHTS

Anotações

Corpus bem anotados e adequados ao domínio são essenciais nas tarefas de Reconhecimento de Entidades Nomeadas.

Códigos acessórios

Soluções deste tipo podem requerer uma série de códigos e serviços acessórios fora das áreas e subáreas da Inteligência Artificial.

Expressões textuais

É possível criar expressões textuais complexas como Entidades Nomeadas a partir de recursos como “*sense2vec*”.

Tratamento dos dados

Para um trabalho mais robusto é necessário preparar melhor o conjunto de dados, criando estruturas que permitam estatísticas automatizadas.

Hiperparâmetros

Seja através do código explícito ou através da parametrização via arquivos de configuração, entender bem os hiperparâmetros será sempre essencial ao modelo.

Do protótipo à produção

Em problemas mais simples, com escopo limitado, buscar soluções com pipelines que permitam tratar o modelo do protótipo à produção pode ser a melhor escolha.



CONCLUSÕES E TRABALHOS FUTUROS

- O modelo criado demonstrou que o conceito proposto é viável, mas ainda serão necessárias melhorias, tais como: maior base de treinamento com anotações mais amplas; ajustes detalhados nos parâmetros da rede neural; melhor preparação e pré-processamento dos dados de entrada.
- O trabalho futuro planejado para o curto prazo será justamente aplicar as melhorias propostas e submeter o protótipo para a Ouvidoria Geral do Estado. No médio prazo deve-se avaliar se é possível adaptar o modelo para generalizar o Reconhecimento de Entidades Nomeadas como dados sensíveis em outros domínios.
- Após o aceite da Ouvidoria Geral do Estado a solução será compartilhada no Github.



OBRIGADO!

Perguntas?

emerson.moura@secont.es.gov.br

emerson.vix@gmail.com