



CÂMARA DOS DEPUTADOS
DIRETORIA DE INOVAÇÃO E TECNOLOGIA DA INFORMAÇÃO
Serviço de Ciência de Dados

Formulário

RESPOSTA A SOLICITAÇÃO DE INFORMAÇÃO - LAI

IDENTIFICAÇÃO DA SOLICITAÇÃO

Protocolo Fale Conosco nº: 201029-000108

Assunto: Uso de tecnologias de Inteligência Artificial pela Câmara

Gostaria de obter informação se o órgão utiliza tecnologias de inteligência artificial, algoritmo ou sistema automatizado de decisão de aprendizado por máquinas em suas iniciativas. Em caso positivo, gostaria de obter as seguintes informações:

Qual o nome da ferramenta utilizada (se houver)?

Ulysses – a Inteligência Artificial (IA) da Câmara

Qual a categoria de aplicação da ferramenta? Por exemplo: classificação de imagens, reconhecimento facial, sistemas de recomendação, estimativa da risco, chatbots..

O Ulysses representa todas as iniciativas de IA da Câmara dos Deputados. Atualmente, há em produção serviços de classificação de textos (tematização de discursos, tematização de legislação, tematização de propostas legislativas, tematização de demandas para Consultoria Legislativa), reconhecimento facial de Deputados, agrupamento de comentários de enquetes por similaridade.

A ferramenta foi desenvolvida interna ou externamente?

Internamente.

Foi comprada pronta? Se sim, de qual fornecedor?

Não se aplica.

Qual o modelo estatístico utilizado na ferramenta? Por exemplo: regressão logística, árvore de decisão, redes neurais.

Os modelos atualmente em produção são:

Tematizações (todos os modelos): FastText (<https://fasttext.cc/>).

Reconhecimento facial de Deputados: redes neurais, utilizando a ferramenta fast.ai (<https://github.com/fastai/fastai>).

Agrupamento de comentários de enquetes por similaridade: redes neurais para representação das sentenças (BERT provido pela ferramenta huggingface.co) e K-means para agrupamento (scikit-learn.org).

Qual o grau de apoio que a ferramenta oferece? Por exemplo: ferramenta faz diagnósticos e

toma decisões, faz diagnósticos e sugere ações, ou faz diagnósticos, mas não sugere ações

De forma geral, a ferramenta não toma decisões no sentido de decisões gerenciais ou estratégicas. A ferramenta apoia tarefas repetitivas e que se aplicam a um volume grande de conteúdos, como textos, imagens etc. A seguir, um detalhamento dessas tarefas.

O serviço de classificação de textos é utilizado em várias aplicações.

- No Portal da Câmara, as páginas temáticas possuem áreas (propostas legislativas, discursos, estudos e legislação) que são alimentadas a partir da tematização feita automaticamente;
- No sistema de cadastro de propostas legislativas, para cada proposta legislativa a ser classificada, sugere-se as três classificações mais prováveis de acordo com o modelo e uma pessoa seleciona as aplicáveis;
- No sistema da Consultoria Legislativa, para cada demanda a ser distribuída para as consultorias temáticas, sugere-se as três áreas mais prováveis de acordo com o modelo e uma pessoa seleciona a área apropriada.

O reconhecimento facial de Deputados é integrado a um aplicativo interno de uso do gabinete parlamentar para mostrar automaticamente as fotos oficiais registradas pela Câmara com aquele Deputado.

Agrupamento de comentários de enquetes por similaridade é utilizado para apoiar a análise humana de milhares de comentários registrados em nossa plataforma de enquetes de propostas legislativas, sem sugerir ações.

Quais as saídas (outputs) da ferramenta? Por exemplo, probabilidade de determinado caso apresentar fraude; ou concessão ou não de crédito.

Nos casos da classificação de textos e reconhecimento facial de Deputados, as saídas do serviço são as probabilidades associadas aos principais resultados (p.ex. 3 temas mais prováveis da proposta legislativa, 3 deputados mais prováveis encontrados em uma imagem).

No caso do agrupamento de comentários de enquetes por similaridade, a saída são os comentários representativos de cada agrupamento e uma lista de comentários mais similares a eles.

Quais são os principais conjuntos de features/variáveis usados como entrada para treinar a ferramenta? Por exemplo: região e IDH de um município para um sistema de estimativa de risco de fraudes; presença no cadastro positivo para concessão de crédito.

Na classificação de textos, a entrada são os textos (p.ex. inteiro teor da proposta legislativa, texto da solicitação de trabalho à Consultoria Legislativa).

No reconhecimento facial de Deputados, a entrada são imagens com faces extraídas dos Deputados.

No agrupamento de comentários de enquetes por similaridade, a entrada são os textos dos comentários das enquetes.

No desenvolvimento e/ou implementação da ferramenta, foram considerados possíveis vieses em seu desempenho? Por exemplo: sua eficácia é melhor com determinados grupos?

Sim, a preocupação com os vieses faz parte do trabalho da equipe. Assim, todos os modelos gerados são catalogados com suas respectivas métricas de desempenho, e nos casos de

problemas de classificação, há registro de *classification report*, detalhando métricas por classes, permitindo verificar classes com desempenho inferior para investigação. Na criação dos modelos, também são aplicadas técnicas como aumento de dados (*data augmentation*) e definição de pesos variáveis de acordo com o tamanho das amostras das classes.

No desenvolvimento e/ou implementação da ferramenta, foi considerada a obrigatoriedade de revisão humana das decisões automatizadas?

Como detalhado em resposta acima, a maior parte das aplicações dos serviços de IA na Câmara não automatizam ações, estando focada no apoio às pessoas de forma que a revisão humana das predições é parte inerente ao processo de trabalho. Nos casos em que há automatização a partir do uso de IA, houve avaliação de gestores sobre os possíveis impactos de erro nos mecanismos de IA e se decidiu que os benefícios gerados compensam os riscos, tendo em vista que as aplicações automatizadas (páginas temáticas do Portal e seleção de fotos para os Deputados) apenas possibilitam outras vias de acesso mais eficiente às informações já disponíveis de outras formas.

Atenciosamente,
Câmara dos Deputados