

Ficha de Curso

Classificação Processual e IA: Parceria entre CNJ, PNUD e Universidade de Fortaleza (UNIFOR)

Curso: Classificação Processual e IA: Parceria entre CNJ, PNUD e Universidade de Fortaleza (UNIFOR)

Modalidade: Ensino a Distância – Autoinstrucional

Público-alvo: Servidores técnicos do Poder Judiciário brasileiro, com atuação na área de Tecnologia da Informação e no desenvolvimento de aplicações para a Plataforma Digital do Poder Judiciário Brasileiro – PDPJ-Br.

Carga horária total: 6 horas

1. Contextualização

O Projeto de Classificação Processual foi idealizado e desenvolvido no âmbito do Programa Justiça 4.0. Ele é fruto de uma parceria entre o CNJ, o PNUD e a Universidade de Fortaleza (UNIFOR), no âmbito do Programa Justiça 4.0.

O Projeto de Classificação Processual classifica automaticamente o assunto de um processo a partir da análise do texto da sua petição inicial. Os assuntos identificados por ele seguem a Tabela Processual Unificada do Conselho Nacional de Justiça (TPUs/CNJ) e abrangem as classes existentes no dataset usado para o seu desenvolvimento. Cabe mencionar que o dataset foi extraído a partir da plataforma CODEX. Foram usadas as petições disponíveis na plataforma naquele momento. Portanto, algumas classes da TPU não estão representadas devido à ausência de dados.

Na sua construção foram aplicadas técnicas de Processamento de Linguagem Natural e algoritmos de Machine Learning. O projeto foi concluído e entregue no início de 2023. Nesse curso reunimos e organizamos vídeos e materiais do projeto para apoiar o estudo dos subprodutos entregues. São eles:

- O **Knowledge Graph Law** (KGL) é um grafo de conhecimento da legislação brasileira. Ele contém as leis brasileiras, seus conteúdos, seus assuntos em consonância com TPU, bem como as conexões entre os diversos mecanismos legais existentes em nosso sistema jurídico. Ele é usado para validar a legislação extraída do texto. É um recurso que pode ser utilizado no desenvolvimento de aplicações de IA para o Poder Judiciário.

- O **Trata_texto** é responsável pelo pré-processamento dos textos jurídicos. Sua contribuição está nas normalizações textuais que realiza. Ele padroniza datas, abreviações, siglas, citações de leis e instrumentos normativos, bem como súmulas e súmulas vinculantes. A normalização de textos melhora o desempenho dos algoritmos de Inteligência Artificial. Este módulo pode ser usado em outras aplicações.

- O **AILA** corresponde ao Trata-texto na forma de plugin, funcionando como uma extensão para o navegador Google Chrome. Ele trabalha em campos de textos do tipo iframe, comumente utilizados em soluções web para a escrita de textos, inclusive no PJe. Seu principal objetivo é auxiliar o usuário durante a digitação apresentando sugestões ao reconhecer legislações, padronizando as citações e oferecendo ao usuário a consulta do texto da lei, de acordo com o Knowledge Graph Laws.

- Modelo **NER**: é um modelo de Inteligência Artificial que reconhece entidades nomeadas. Neste projeto ele está sendo usado para reconhecer apenas entidades relacionadas à legislação (principal e/ou acessória), mas seu uso pode ser ampliado.

As entidades extraídas por esse modelo são validadas no grafo de conhecimento (KGL).

- Modelo **Classificador de Assuntos**: é um modelo de Inteligência Artificial baseado em Redes Neurais que classifica os textos pré-processados nos assuntos da TPU. Na arquitetura desse classificador foi usado o modelo de linguagem BERTimbau.

O projeto e seus subprodutos estão disponíveis na plataforma SINAPSES.

2. Objetivos

O objetivo da presente capacitação é fornecer subsídios e exemplos para que o/a cursista possa entender, reproduzir e aplicar as soluções de IA disponibilizadas no âmbito do Projeto de Classificação Processual. É objetivo também habilitar o/a cursista a entender conceitualmente a arquitetura dos modelos de Inteligência Artificial (IA) construídos no âmbito do projeto.

Nesse sentido, ao final do curso, o/a participante deverá ser capaz de entender a arquitetura do classificador, especialmente no que se refere ao papel de cada subproduto mencionado no fluxo de execução do classificador. Espera-se ainda que o/a participante, visando o aprimoramento do funcionamento do seu tribunal, consiga reproduzir soluções similares que possam ser utilizadas no âmbito do seu tribunal.

3. Programa

O curso está dividido em **6 Módulos**, distribuídos da seguinte maneira:

Módulo 1: Grafo de Conhecimento da Justiça Brasileira

- Grafo de Conhecimento da Justiça Brasileira – Arquitetura Geral
- Grafo de Conhecimento da Justiça Brasileira – Detalhe Técnicos
- Grafo de Conhecimento da Justiça Brasileira – Aplicações

Módulo 2: Classificador de Assuntos – Parte 1

- Classificador de Assuntos – Construção do Dataset
- Classificador de Assuntos – Definição a partir das TPUs
- Classificador de Assuntos – Modelos de Classificação

Módulo 3: Classificador de Assuntos – Parte 2

- Classificador de Assuntos – Implementação dos Modelos de Classificação I
- Classificador de Assuntos – Implementação dos Modelos de Classificação II
- Classificador de Assuntos – Implementação dos Modelos de Classificação III
- Classificador de Assuntos – Implementação dos Modelos de Classificação IV

Módulo 4: Trata Texto em Lote

- Trata Texto em Lote I
- Trata Texto em Lote II

Módulo 5: Artificial Intelligence Law Assistant - AILA

- AILA I
- AILA II

Módulo 6: Modelo NER - Reconhecimento de Entidades Nomeadas

- Modelo NER

Nota: o curso sofreu refinamentos durante seu processo de finalização, de forma que pode ocorrer menções de nomenclaturas que foram descontinuadas, como Projeto ALICIA, por exemplo. Por gentileza, desconsiderar essas menções.

4. Procedimentos didáticos

Como suporte ao Ensino Remoto será utilizada a infraestrutura do Ambiente Virtual de Aprendizagem (AVA) da plataforma moodle do Centro de Formação e Aperfeiçoamento de Servidores do Poder Judiciário (CEAJUD). Na Plataforma estarão disponíveis os seguintes materiais:

- Ficha do Curso;
- Links para acesso às videoaulas legendadas (módulos gravados);
- Material Complementar em formato de slides;
- Atividade avaliativa de desempenho para certificação; e
- Avaliação sobre a qualidade do curso.

Destaca-se que o presente curso, no modelo autoinstrucional nesta plataforma, é originário da versão síncrona e online do treinamento previamente ofertado. Nesse sentido, eventuais cortes no áudio ou na imagem dos conteudistas durante as videoaulas podem ser fruto de instabilidade da rede no momento da gravação.

5. Carga horária do(a) cursista

A seguir, é apresentada a carga horária total do curso.

Atividades	Carga horária (horas)	Fração (%)
Estudo individual de conteúdo pré-aula	0h30	8,33%
Estudo individual em aula assíncrona	3h	50%
Estudo individual pós-aula	1h30	25%

Atividade Avaliativa e Avaliação do Curso	1h	16,67%
Total	6 horas	100%

6. Avaliação de desempenho

O desempenho no curso será determinado por nota na atividade avaliativa realizada ao final do curso e pelo preenchimento da avaliação sobre a qualidade do mesmo. Requer-se, no mínimo, 70% de aproveitamento na atividade avaliativa para obtenção de certificado.

7. Referências bibliográficas

FACELI, Katti; LORENA, Ana Carolina; GAMA, João; CARVALHO, André Carlos Ponce de Leon Ferreira de. Inteligência artificial: uma abordagem de aprendizado de máquina. [S.l: s.n.], 2021.

RUSSEL, S.; NORVIG, P.; Artificial Intelligence: A Modern Approach. Prentice-Hall, Second Edition, 2013.

8. Requisitos Técnicos

Para cursar a presente capacitação, faz-se necessário possuir:

- Computador e/ou tablet capazes de acessar a plataforma moodle do CEAJUD;
- Versões mais atualizadas de Windows, Linux ou Mac;
- Conexão de Internet.