



TECH CHALLENGE FASE 2

TECH CHALLENGE

Tech Challenge é o projeto que engloba os conhecimentos obtidos em todas as disciplinas da fase. Esta é uma atividade que, em princípio, deve ser desenvolvida em grupo. É importante atentar-se ao prazo de entrega, uma vez que essa atividade é obrigatória, valendo 90% da nota de todas as disciplinas da fase.

DESAFIO

Após o sucesso no desenvolvimento de modelos de machine learning para o diagnóstico médico (Fase 1), o hospital universitário agora enfrenta dois desafios importantes que podem ser solucionados com as técnicas de algoritmos genéticos e processamento de linguagem natural.

Os alunos **podem escolher um dos seguintes projetos:**

Projeto 1: Otimização de Modelos de Diagnóstico

O hospital precisa melhorar a precisão e eficiência dos modelos de diagnóstico desenvolvidos na Fase 1. O desafio é utilizar algoritmos genéticos para otimizar os hiperparâmetros desses modelos, além de incorporar capacidades iniciais de processamento de linguagem natural por meio de LLMs para melhorar a interpretabilidade dos resultados para os profissionais de saúde.

Projeto 2: Otimização de Rotas para Distribuição de Medicamentos e Insumos

O sistema hospitalar enfrenta desafios logísticos na distribuição eficiente de medicamentos e insumos entre suas diversas unidades e para atendimento domiciliar. O desafio é desenvolver um sistema de otimização de rotas utilizando algoritmos genéticos para resolver esse problema de "caixeiro viajante médico", além de utilizar LLMs para gerar relatórios e instruções claras para as equipes de entrega.

Ambos os projetos serão fundamentais para preparar a infraestrutura necessária para o assistente médico mais avançado, previsto na Fase 3.

OBJETIVO

Desenvolver uma solução para otimização de modelos de ML médicos existentes, além de implementar recursos iniciais de processamento de linguagem natural para melhorar a interpretação e apresentação dos diagnósticos. **Alternativamente**, os alunos podem optar por desenvolver um sistema de otimização de rotas para entrega de medicamentos e insumos médicos utilizando algoritmos genéticos.

REQUISITOS OBRIGATÓRIOS

Entregas técnicas para o PROJETO 1 (Otimização de Modelos de Diagnóstico):

1. Otimização via Algoritmos Genéticos

- Implementar um algoritmo genético para otimização de hiperparâmetros dos modelos de diagnóstico desenvolvidos na Fase 1:
 - Definir uma codificação adequada (representação de genes) para os hiperparâmetros relevantes;
 - Implementar operadores de seleção, cruzamento e mutação;

- Definir uma função fitness baseada nas métricas de desempenho dos modelos (accuracy, recall, F1-score, etc.).
- Comparar o desempenho dos modelos otimizados com os modelos originais;
- Realizar ao menos três (3) experimentos com diferentes configurações do algoritmo genético (tamanho da população, taxas de mutação, etc.).

Configurar recursos de escalabilidade automática para lidar com variações de demanda

- Implementar monitoramento e logging adequados para tracking de desempenho;
- Documentar arquitetura e decisões de implementação.

Observação: a implementação em nuvem é opcional e pode ser considerada para pontuação extra.

3. Integração com LLMs para Interpretação de Resultados

- Integrar uma LLM pré-treinada (GPT, Falcon, LLaMA, etc.) para:
 - Gerar explicações em linguagem natural dos diagnósticos produzidos pelos modelos;
 - Transformar dados numéricos e estatísticos em insights acionáveis para médicos;
 - Preparar a base para a futura integração com dados textuais na Fase 3;
- Implementar técnicas de prompt engineering para obter respostas relevantes e adequadas ao contexto médico;
- Avaliar a qualidade das interpretações geradas.

Entregas técnicas para o PROJETO 2 (Otimização de Rotas Médicas):

1. Sistema de Otimização de Rotas via Algoritmos Genéticos

- Desenvolver um sistema que resolva o problema do caixeiro viajante (TSP) para otimizar rotas de entrega de medicamentos e insumos:
 - Implementar a representação genética adequada para rotas;
 - Desenvolver operadores genéticos especializados (seleção, crossover, mutação) para o problema de roteamento;
 - Criar uma função fitness que considere distância, prioridade de entregas e outras restrições relevantes.
- Incluir restrições realistas como:
 - Prioridades diferentes para entregas (medicamentos críticos vs. insumos regulares);
 - Capacidade limitada de carga dos veículos;
 - Autonomia limitada dos veículos (distância máxima que pode ser percorrida);
 - Múltiplos veículos disponíveis (ampliando para o problema de roteamento de veículos - VRP);
 - Outras restrições que achar interessante.
- Visualizar as rotas otimizadas em um mapa para fácil interpretação.

Importante: como ponto de partida, será fornecido um código base de TSP em Python. Os alunos deverão modificar esse código para incorporar as restrições adicionais e torná-lo mais realista para o contexto hospitalar.

3. Integração com LLMs para Geração de Instruções e Relatórios

- Utilizar uma LLM pré-treinada para:
 - Gerar instruções detalhadas para motoristas e equipes de entrega com base nas rotas otimizadas;
 - Criar relatórios diários/semanais sobre eficiência de rotas, economia de tempo e recursos;
 - Sugerir melhorias no processo com base nos padrões identificados.

- Implementar prompts eficientes para extrair informações úteis da LLM;
- Permitir que o sistema responda as perguntas em linguagem natural sobre as rotas e entregas.

4. Código e Organização (para ambos os projetos)

- Projeto Python bem estruturado, utilizando ambiente virtual (Poetry, Pipenv ou venv);
- Documentação detalhada, incluindo diagramas de arquitetura;
- Testes automatizados para validação de funcionalidades;
- Se optou pela implementação em nuvem: infraestrutura como código (IaC) para provisionamento dos recursos.

ENTREGÁVEIS DA FASE 2

Repositório Git:

- Código-fonte completo;
- Documentação da API (se aplicável);
- Scripts ou notebooks de demonstração;
- Se optou pela implementação em nuvem: arquivos de configuração para implantação.

Relatório técnico explicando:

Projeto 1 (Otimização de Modelos):

- Implementação do algoritmo genético e resultados da otimização de hiperparâmetros;
- Integração com LLMs: abordagem, prompts utilizados e avaliação da qualidade;
- Comparativo de desempenho entre os modelos originais e otimizados;

- Desafios enfrentados e soluções implementadas;
- Se optou pela implementação em nuvem: Arquitetura da Solução.

Projeto 2 (Otimização de Rotas):

- Implementação do algoritmo genético para roteamento a partir do código base fornecido;
- Estratégias para lidar com restrições adicionais (prioridades, capacidade de carga, autonomia dos veículos, múltiplos veículos);
- Integração com LLMs para geração de instruções e relatórios;
- Comparativo de desempenho com outras abordagens de roteamento;
- Visualizações e análises das rotas otimizadas;
- Se optou pela implementação em nuvem: Arquitetura da Solução.

Vídeo de demonstração:

- Realizar o upload no YouTube ou Vimeo (utilizar as configurações “público” ou “não listado”). A duração do vídeo deve ser de até 15 minutos;
- Demonstração do sistema em execução;
- Explicação dos diferentes componentes da solução;
- Apresentação dos resultados da otimização via algoritmos genéticos;
- Demonstração da integração com LLMs.



POSTECH