

Atividade 01

INF 2137 - Engenharia de software para ciência de dados

Pedro Henrique Cardoso Paulo

pedrorjpaulo.phcp@gmail.com

Professor: Marcos Kalinowski



Departamento de Engenharia Mecânica
PUC-RJ Pontifícia Universidade Católica do Rio de Janeiro
abril, 2023

Atividade 01

INF 2137 - Engenharia de software para ciência de dados

Pedro Henrique Cardoso Paulo

abril, 2023

1 Introdução

Esta tarefa tem por objetivo exercitar as competências de especificação de softwares inteligentes usando a metodologia do PerSpecML.

2 Enunciado

2.1 Contexto

Uma empresa fornece uma solução de marketplace em que diferentes anunciantes (sellers) podem anunciar seus produtos para venda. Como diferentes tipos de produtos requerem diferentes informações cadastrais (por exemplo, camisas possuem informações como material, tamanho e cor, enquanto telefones celulares possuem informações como processador, espaço de armazenamento, capacidade da bateria, entre outros), a empresa deseja implementar uma classificação automática dos tipos de produtos com base nos títulos dos anúncios utilizando aprendizado de máquina. Desta forma, ela será capaz de apresentar para o anunciante os formulários apropriados para o cadastro das informações detalhadas, maximizando as chances de venda.

A empresa acredita que com essa classificação automática e a consequente melhoria das informações cadastrais ela será capaz de aumentar as vendas do marketplace em 5%. Os usuários esperam poder realizar o cadastro de forma prática, sem complicações de classificação manual. Do ponto de vista mais técnico, espera-se que o modelo de aprendizado de máquina tenha uma acurácia de pelo menos 80% na classificação. Além disso, devido ao grande volume de anunciantes, o modelo deverá ser capaz de processar pelo menos 10.000 requisições por minuto.

2.2 User Story

Como anunciante do marketplace **quero** obter um formulário com as informações cadastrais corretas a partir do título do meu produto **para** poder cadastrar o produto de forma precisa, maximizando as chances de venda.

2.3 Tarefa

Realize a especificação utilizando o template da técnica PerSpecML do Miro disponível neste [link](#). Como resolução da tarefa no ambiente, poste um PDF com um link para a sua cópia do board do Miro. Opcionalmente, inclua no PDF qualquer outra informação que julgar pertinente, como, por exemplo, feedback sobre a aplicação da técnica.

3 Solução

3.1 Link

- [Link para o Miro com a resolução](#)

3.2 Premissas adotadas e pontos relevantes

- Como não houve requisito explícito de tempo máximo para inferência, foi adotado 5 segundos. O requisito de 10000 operações por minuto foi explicitado também, mas este pode ser compensado por meio da paralelização do serviço em múltiplos containeres.

- Foi entendido pela explicação da tarefa que o sistema seria *forceful* e que seria mais danoso para o usuário ter uma predição errada do que uma não predição. Dessa forma, isso foi explicitado no quadro.
- Para garantir algum nível de *feedback* dos cadastradores, assumiu-se que eles teriam uma opção de ignorar a saída do sistema e ir para um formulário genérico. Pelos motivos explicitados na explicação (muitas categorias), eles não poderiam reclassificar manualmente.
- Como o marketplace fica aberto à compradores, assumiu-se que compradores poderiam dar o feedback sobre a classificação feita.
- Como não sabemos se o marketplace contrata contêineres na nuvem ou usa servidores próprios, ambas as opções foram explicitadas.

3.3 Impressões gerais sobre a técnica

- A aplicação da técnica foi bem bacana pois dá espaço para assuntos bem focados na especificação para softwares inteligentes. Normalmente quando se usa dinâmicas mais genéricas como um *lean canvas* simples, o espaço para divergências é muito grande. Para fases em que já se definiu que o produto será uma solução digital focada em ML, a dinâmica é muito boa.
- A separação dos tópicos parece ajudar especialistas de cada área a se acharem e contribuir. Pessoalmente, ficou fácil achar os tópicos que eu tinha mais domínio
- Não parece ser uma dinâmica fácil de fazer se não tiver um bom grupo envolvido. Pelo próprio exercício, em que fazemos sozinhos, fica evidente que falta a presença de mais gente que conheça do problema para detalhar algumas áreas.