

IA Classificação Mercadológica

Objetivo principal

API Django que classifica produtos em uma estrutura mercadológica (departamento > categoria > subcategoria) usando IA generativa. Recebe descrição/EAN/NCM de um produto, consulta a base local antes de acionar a IA (evitando reprocessamento) e cadastra automaticamente o produto quando classificado com sucesso. Suporta classificação unitária e em lote, registra telemetria de uso (tokens e quantidade de produtos por empresa/grupo) e gera/envia mensalmente uma planilha de bilhetagem por e-mail.

Principais tecnologias

Python / Django 5 / Django REST Framework / LangChain (integração plugável com provedores de LLM) / PostgreSQL (pgvector) / Docker

Principais funcionalidades

- Classificação unitária de produtos: recebe descrição/EAN/NCM e retorna departamento, categoria e subcategoria, usando a base existente ou acionando a IA quando necessário
- Classificação em lote (v1 e v2): permite enviar múltiplos produtos para classificação, com modelo v2 baseado em lotes assíncronos com status (criado, pendente, processando, concluído, falhou) e callback de retorno
- Cadastro automático de produtos: quando um produto com EAN é classificado com sucesso, é inserido na base para consultas futuras sem nova chamada à IA
- Votação de classificação: endpoint para aprovar/reprovar a classificação de um produto, com comentário obrigatório em caso de reprovação
- Normalização de descrições de produtos: upsert em lote de descrições normalizadas por GTIN/NCM, usadas para enriquecer o prompt enviado à IA
- Configuração dinâmica de modelo de IA: cadastro via admin do modelo ativo (biblioteca, parâmetro de chave de API, prompts unitário e de lote), com instalação automática do pacote pip necessário
- Telemetria de uso de IA: registro de tokens consumidos, tipo de operação (recuperação em base, unitário, em massa) e quantidade de produtos por usuário/grupo (empresa)
- Bilhetagem mensal: comando agendado (cron) que gera planilha XLSX de consumo por empresa e envia por e-mail via SMTP
- Multiempresa via grupos: cadastro de empresas/usuários com token de API, associados a grupos que segmentam a telemetria por ambiente/empresa
- Importação de classificações: comando de management para carregar a estrutura de departamentos/categorias/subcategorias a partir de um arquivo JSON

Principais integrações

- Provedores de LLM via LangChain: modelo de IA plugável e configurável em runtime (ex: OpenAI, Google), definido por classe/pacote cadastrados no admin
- PostgreSQL com pgvector: banco de dados principal, preparado para buscas vetoriais
- SMTP: envio automático da planilha mensal de bilhetagem de uso da IA por e-mail
- API de cadastro de empresas/usuários: endpoint de integração para provisionamento automático de empresas e emissão de token de acesso
- Callback HTTP (api_callback_url): notificação de sistemas externos sobre a conclusão de um lote de classificação

Revision #1

Created 18 August 2026 18:29:49 by Pedro Souza

Updated 18 August 2026 18:29:49 by Pedro Souza