



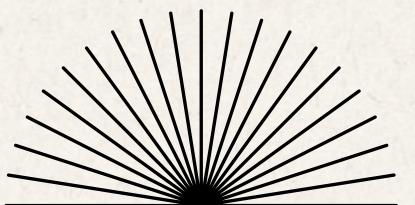
Grupo Not a Name | 11/2025

LLM MATCHMAKER

O melhor modelo para o melhor resultado

NOME DO PROJETO:
LLM MatchMaker

APRESENTADO POR:
Grupo NaN



Nosso time



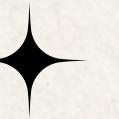
Erik Martins

RM554695



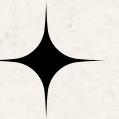
Gustavo Borges

RM554612



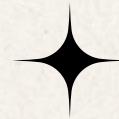
Marco Antonio

RM557859



Regiane Patrícia

RM556851



Agenda

03	Contexto
04	Objetivos
05	O problema
06	A solução
07	Impactos da solução
08	Benefícios esperados
09	Público-Alvo
10	Arquitetura
11	O modelo
12	Protótipo
13	Conclusão

01 O número de LLMs disponíveis aumentou drasticamente, cada um com níveis distintos de precisão, velocidade, custo e privacidade.

02 Usuários não especialistas enfrentam dificuldade em entender qual modelo atende melhor suas necessidades específicas.

03 Essa escolha equivocada pode gerar respostas inadequadas, custos maiores e uma experiência de uso ineficiente.



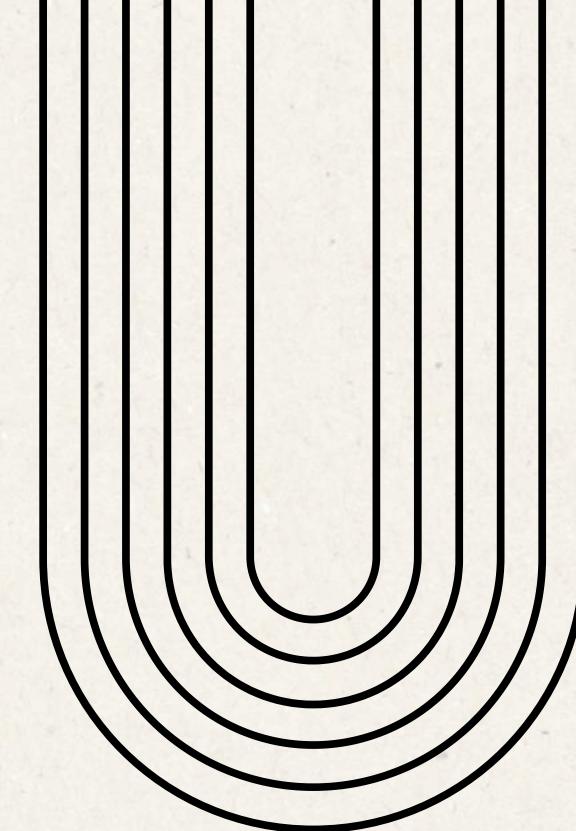
03/13

Contexto

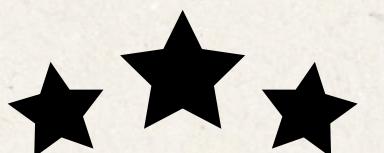
O ecossistema de modelos de linguagem cresceu rapidamente, trazendo grande diversidade de capacidades, custos e limitações.

Para usuários e equipes, escolher o modelo adequado para cada tarefa se tornou um desafio complexo.

Objetivos



Definir claramente o propósito central do projeto e o que buscamos alcançar com a PoC.



Recomendar automaticamente o melhor LLM para cada tarefa.

Oferecer uma escolha inteligente baseada em contexto, restrições e intenção do usuário.



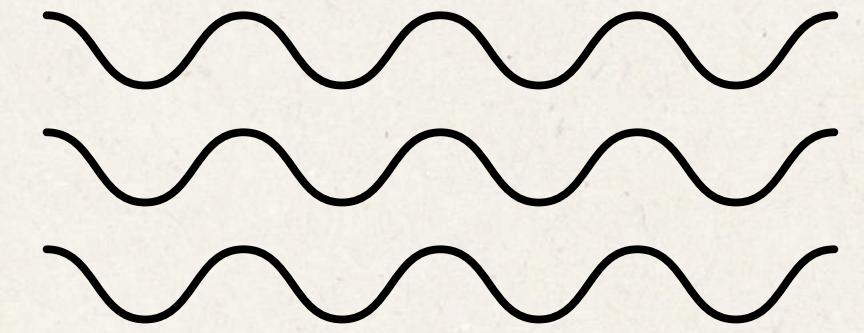
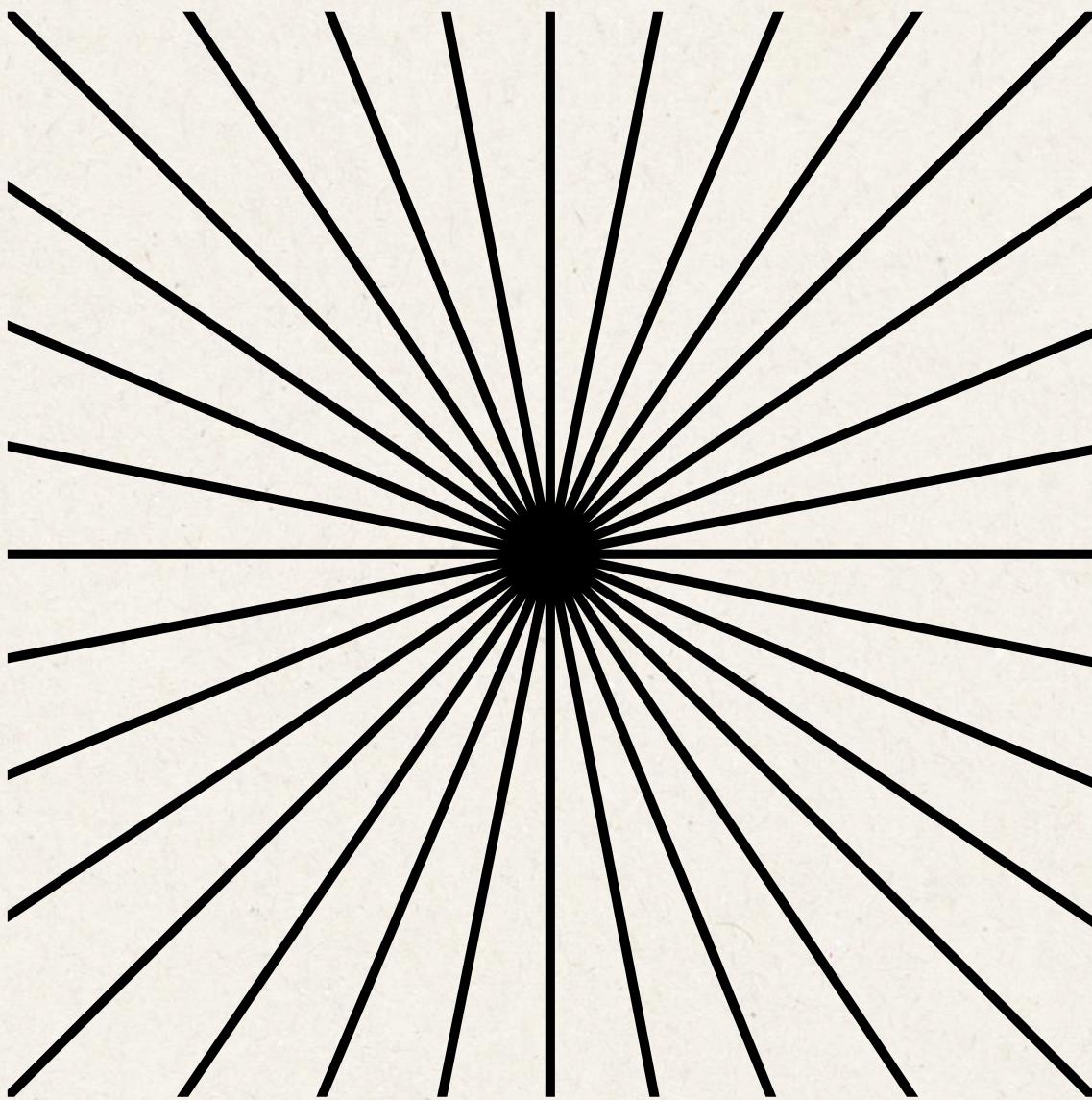
Simplificar a experiência de quem não é especialista.

Ajudar usuários comuns a navegar a complexidade dos modelos de linguagem com clareza e segurança.



Validar um modelo de ML funcional e integrado a uma API.

Demonstrar uma PoC completa: classificação → API (FastAPI) → integração via MCP.



O Problema

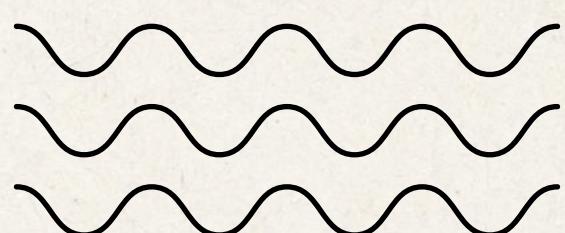
Com a rápida expansão dos modelos de linguagem, escolher qual LLM usar tornou-se uma tarefa confusa e técnica. Cada modelo possui custos, latência, precisão, limitações e requisitos diferentes – e usuários não especialistas não conseguem tomar essa decisão com segurança. A escolha incorreta gera respostas ruins, desperdício de recursos e experiências inconsistentes.

Solução Proposta

LLM Matchmaker

Um sistema inteligente que recomenda automaticamente o melhor modelo de linguagem para cada tarefa.

Ele interpreta a descrição do usuário, extrai as principais características e identifica qual LLM tem maior probabilidade de entregar o melhor desempenho no cenário apresentado.



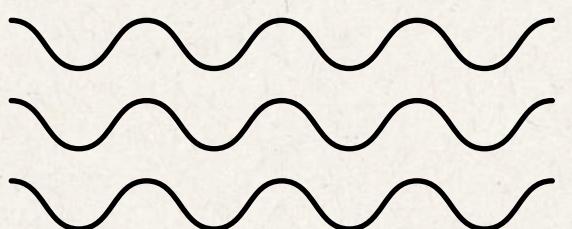
Como funciona

O usuário fornece tarefa, contexto e restrições. O sistema transforma isso em features, utiliza um classificador treinado e retorna o modelo ideal com justificativa e score de confiança – acessível via API (FastAPI) e integrado a agentes por MCP.

Impactos da Solução

A introdução de um recomendador inteligente de LLMs possibilita decisões mais seguras, rápidas e assertivas no uso de modelos de linguagem.

A solução reduz a complexidade para o usuário final, melhora a utilização de recursos computacionais e aumenta a qualidade dos resultados entregues por agentes e aplicações que dependem de IA. Ao automatizar a escolha do modelo adequado, cria-se um fluxo mais eficiente, confiável e alinhado às necessidades reais de cada tarefa.

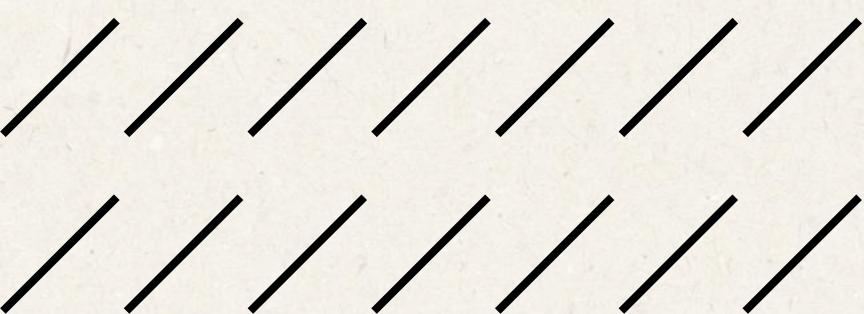


Benefícios Esperados

O LLM Matchmaker traz ganhos diretos, como maior precisão nas recomendações, redução de tentativas e erros na seleção de LLMs e uso mais eficiente das ferramentas disponíveis.

Do ponto de vista operacional, a solução diminui o tempo gasto em análises manuais e facilita a adoção de IA por usuários não especialistas.

Além disso, a integração com agentes modernos permite experiências mais intuitivas e automáticas, ampliando o potencial de uso em diversos contextos.

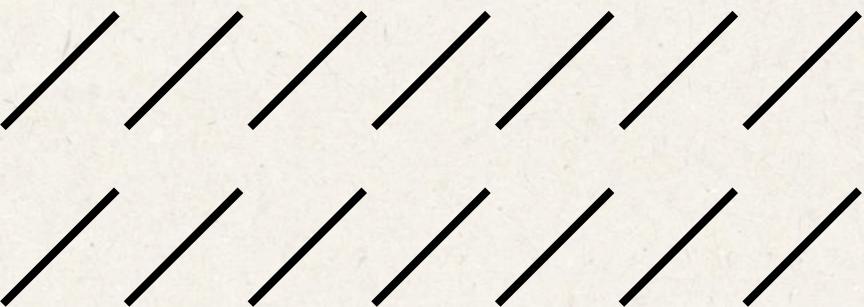
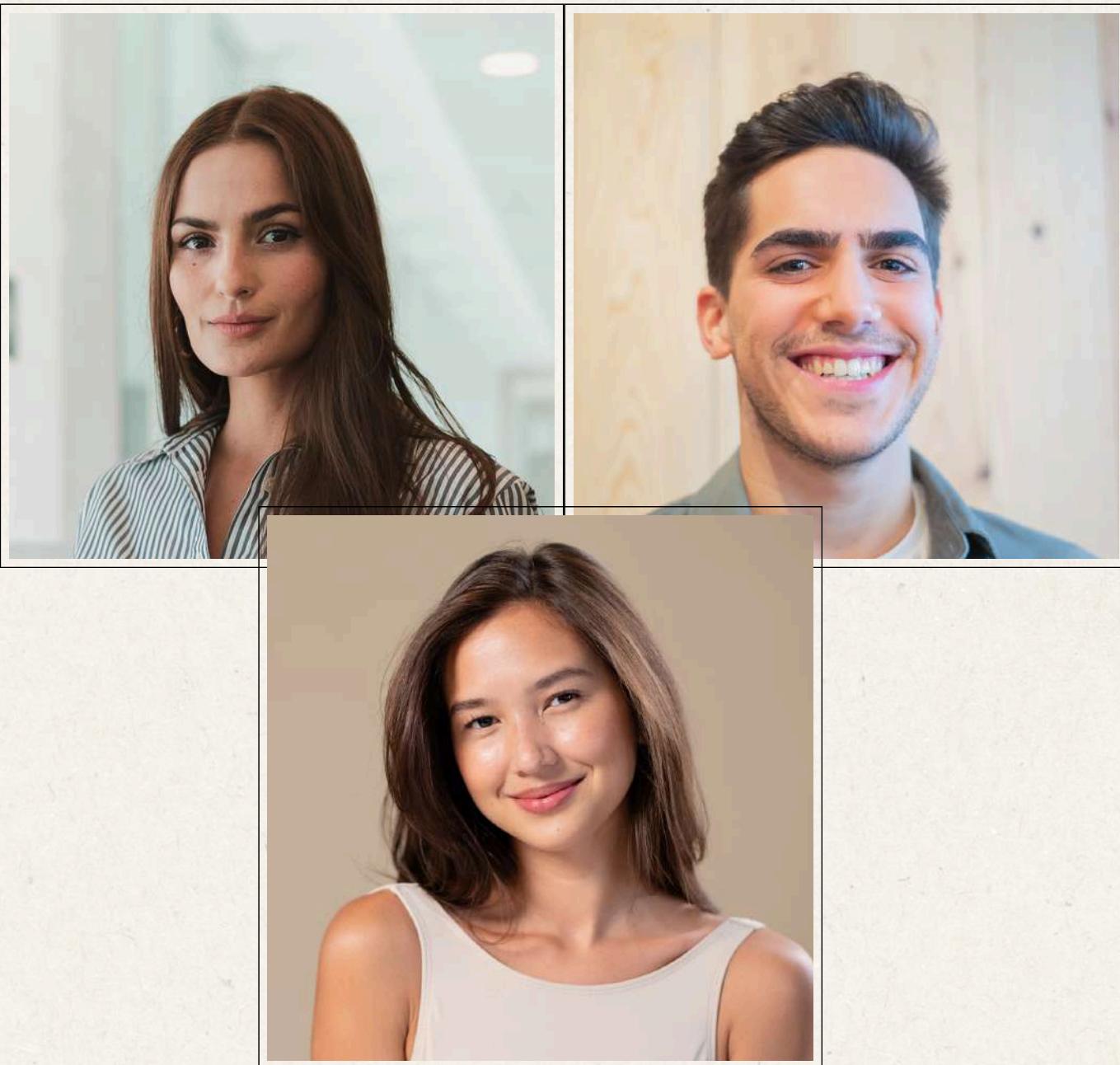


Público-Alvo

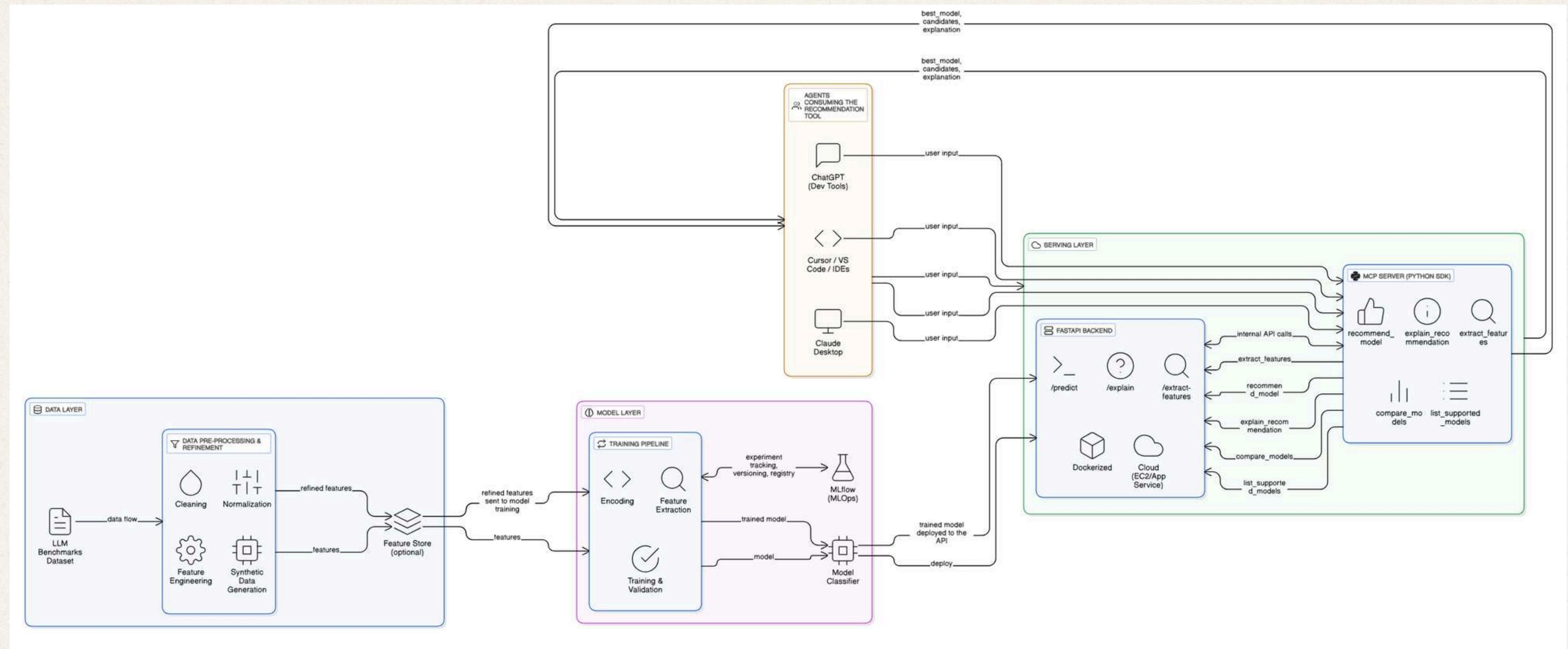
A solução é voltada para pessoas e equipes que utilizam IA no dia a dia, mas enfrentam desafios na escolha entre diferentes modelos de linguagem.

Inclui usuários não técnicos que precisam de orientação clara, profissionais que trabalham com múltiplos LLMs em seus fluxos de trabalho e desenvolvedores que desejam integrar recomendações automáticas em agentes, ferramentas e aplicações.

O objetivo é apoiar qualquer pessoa que precise tomar decisões informadas no ecossistema crescente de LLMs.



Arquitetura

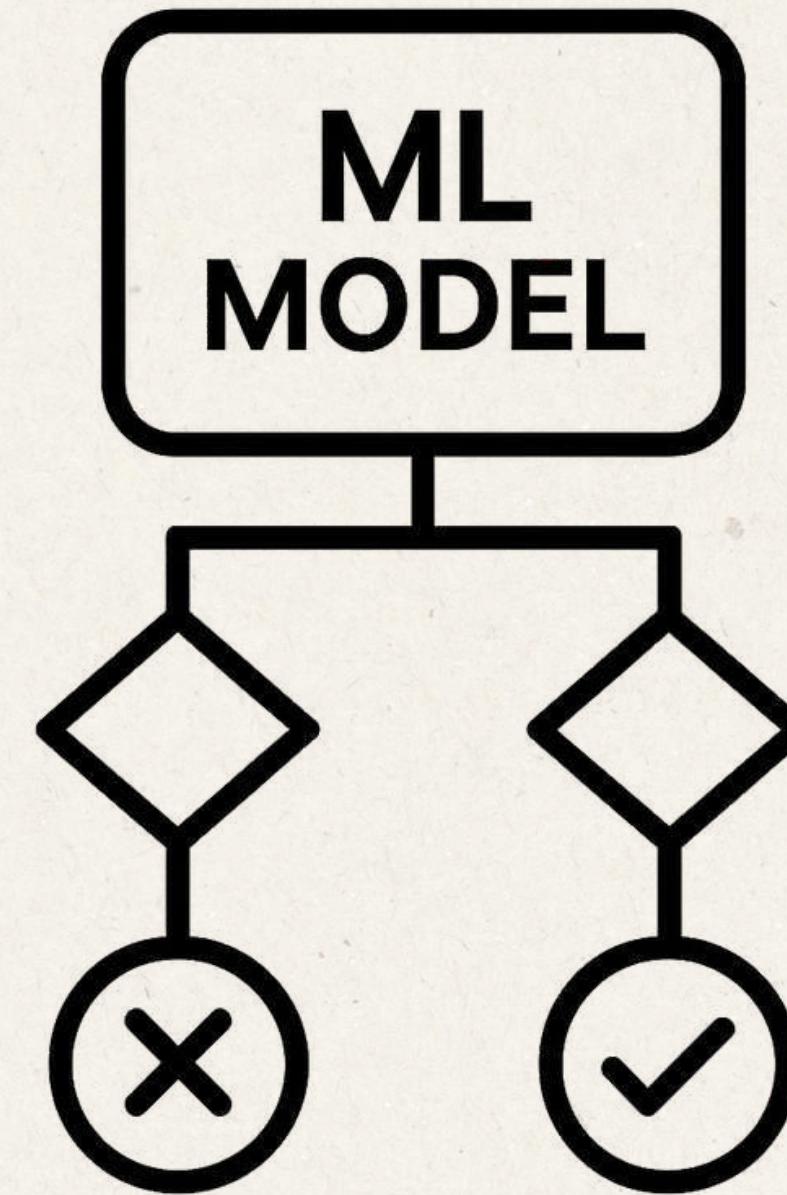


O modelo de ML

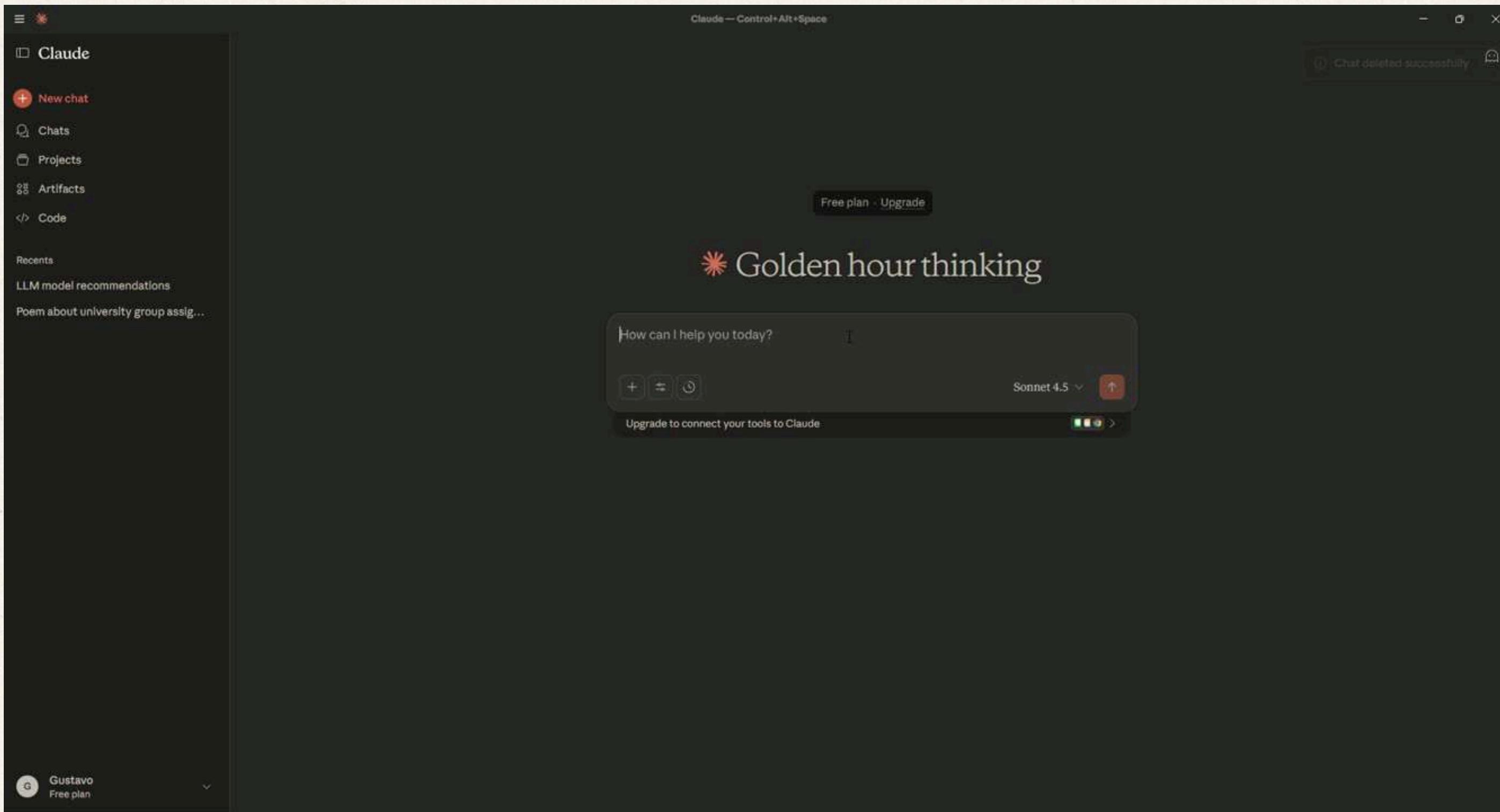
O núcleo do LLM Matchmaker é um modelo de classificação, treinado a partir de dados de benchmarks que descrevem tipos de tarefa, domínios, idioma, requisitos de privacidade, limitações de hardware e preferências como determinismo ou tolerância a alucinações.

Com base nessas informações estruturadas, o modelo prevê qual LLM tem maior probabilidade de entregar o melhor desempenho em cada cenário, oferecendo a recomendação principal, níveis de confiança e alternativas possíveis.

Apesar de simples, o modelo demonstra como uma abordagem de ML pode automatizar a escolha entre diferentes LLMs em casos reais de uso.



Protótipo



Conclusão

O LLM Matchmaker demonstra como uma abordagem simples, mas bem direcionada, pode apoiar a escolha entre diferentes modelos de linguagem de forma mais clara, objetiva e automatizada.

A PoC valida o potencial de integrar um modelo de classificação com uma API moderna e um MCP Server, criando uma base sólida para agentes e aplicações utilizarem recomendações inteligentes em tempo real. Embora inicial, a solução abre caminho para evoluções importantes, mostrando como decisões orientadas por contexto podem melhorar o uso de LLMs no dia a dia.



Obrigado

 Pitch

FALE CONOSCO

Representante Gustavo Borges

E-mail RM554612@fiap.com.br