

Trecho	Código	Categoria
assim, da forma como está eu realizo (segue o workflow), no geral nessa ordem,	O desenvolvimento segue o workflow.	Limitação [Geral]
eu ia comentar apenas da questão do ciclo, porque é quando você vai avançando (nas etapas) que você vê a necessidade de voltar [...]	Ciclos de desenvolvimento existem devido a incerteza de correteude	Desafio [Geral]
De forma geral é isso mesmo. E eu acho que passo mais um tempo tratando os dados, atributos.	Processamento de dados como etapa que toma mais tempo	Desafio [Processamento de dados]
	Engenharia de atributos como etapa que toma mais tempo	Desafio [Engenharia de atributos]
No projeto que eu trabalho, os requisitos já estão prontos, são entregues para mim.	Desenvolvedor não participa da especificação de requisitos.	Limitação [Requisitos do modelo]
Sobre a implantação, eu encontrei mais dificuldade, uma coisa é você trabalhar com arquivo csv e gerar um modelo, testar e tudo, outra coisa é implantar ele em um sistema que já existe e que muitas coisas não podem ser alteradas.	A empresa não desenvolve MLES	
	Implantação do modelo em um sistema pré-existente	Limitação [Implantação do modelo]
	Dificuldade de implantar o modelo devido a regras de negócio do sistema pré existente.	Desafio [Implantação do modelo]
Uma outra coisa sobre a implementação é que eu vinha trabalhando com sklearn, e eu sinto que é uma biblioteca mais para você estudar, desenvolver modelo do que para aplicar no mundo real (aplicações reais). Eu vi que existiam outras alternativas mais profissionais, mas tive que me resguardar a esta. [...] Outra alternativa seria o Spark, algumas outras alternativas que são mais escaláveis.	Uso de Frameworks para criação de modelos	Método [Treino do modelo]
	Frameworks utilizados são pouco escaláveis para aplicações reais	Desafio [Treino do modelo]
	Frameworks mais profisisonais existem.	Limitação [Treino do modelo]
apenas um desenvolvedor mesmo, só eu que trabalhava com isso.	Único desenvolvedor em todas as etapas.	Limitação [Geral]
	Presença do desenvolvedor	Ator [Geral]
não, o projeto era bem livre, não tinha apego a nenhum tipo de metodologia, acho que poderia até ajudar, mas ...	Metodologias de desenvolvimento não são seguidas.	Limitação [Geral]
não, também não. Tem que usar o bom senso aí.	Padrões de código não são seguidos.	Limitação [Geral]
essas etapas de treino e avaliação tem que se repetir né? você tá o tempo todo tentando tunar o seu modelo. [...] Isso é o principal.	Ciclos de desenvolvimento existem devido a incerteza de correteude	Desafio [Geral]
nessas quatro do meio (processamento de dados, engenharia de atributos, treino do modelo e avaliação do modelo).	Entrevistado considera ter experiência nas etapas de construção do modelo.	Limitação [Geral]
pessoa? Tem alguém acima de mim (gerente) que avaliava o modelo, os problemas que eu tinha.	Gerente de projeto avalia a qualidade do modelo de acordo com as necessidades do cliente.	Atividade [Avaliação do modelo]
esse modelo ele é não supervisionado, modelo de cluster então eu usava umas métricas de homogeneidade, métricas de completude e teve um momento que eu avalei o silhouette. São métricas um pouco diferente da aprendizagem supervisionada.	Presença do gerente	Ator [Avaliação do modelo]
	Uso de aprendizagem não supervisionada	Atividade [Treino do modelo]
	Criação de modelos para clusterização.	Atividade [Treino do modelo]
	Métricas clássicas para avaliar o modelo.	Atividade [Avaliação do modelo]
assim, eu não tinha a ideia de valor satisfatório, o que eu tentava fazer era maximizar essas métricas, até porque a interpretação delas não é muito simples.	Métricas não possuem valor alvo definido influenciado pelo tipo de aprendizagem.	Limitação [Avaliação do modelo]
	Busca pela maximização das métricas.	Atividade [Avaliação do modelo]
	Métricas para modelos não supervisionados não possuem interpretação simples.	Desafio [Avaliação do modelo]
Outra coisa que fiz foi gerar gráficos, que ilustram um aspecto importante do nosso resultado.	Uso de gráficos para a avaliação do modelo.	Atividade [Avaliação do modelo]

eu colocaria a etapa de treino do modelo, porque muitas vezes para você desenvolver e avaliar um modelo é complicado você testar o que você tá fazendo.	Treino do Modelo como etapa mais problemática.	Desafio [Treino do modelo]
Primeiro porque tem pouco material sobre o desenvolvimento de modelos assim para fins práticos,	Falta de material relacionado a modelos não supervisionado com fins práticos.	Desafio [Geral]
outra coisa é porque é difícil de você testar o código porque muitos modelos por exemplo tem características aleatórias. Então o desenvolvimento não é sentar e fazer.	Característica aleatória do modelo dificulta a verificação.	Desafio [Treino do modelo]
Muitas vezes você também demora para ter seu resultado, então você quer avaliar o seu modelo e leva alguns minutos, isso tudo pode quebrar muito o fluxo do seu trabalho.	Demora para o treino do modelo quebra no ritmo de trabalho.	Desafio [Treino do modelo]
Eu também tive problema na implantação, porque eu tava mais habituado com o treinamento do modelo, avaliação e na hora de implantar eu tive que para um pouco para aprender outras tecnologias, como integrar no sistema em si.	Dificuldade na implantação do modelo devido a necessidade de dominar outras tecnologias.	Desafio [Implantação do modelo]
primeiro o tempo, as vezes demora muito para treinar.	Demora para o treino do modelo.	Desafio [Treino do modelo]
Às vezes você tem algum tipo de problema, e ou a documentação é escassa ou a comunidade não resolveu aquele problema.	Frameworks para aprendizagem não supervisionada com documentação escassa.	Desafio [Treino do modelo]
	Frameworks para aprendizagem não supervisionada com problemas não solucionados	Desafio [Treino do modelo]
Eu diria que (a aprendizagem não supervisionada) não está tão madura assim.	Comunidade não supervisionada está pouco madura.	Desafio [Geral]
dos problemas que eu citei, a questão do tempo você vê enquanto está usando [...]. Os erros que falei, por muitas vezes são erros na execução do algoritmo, então é erro de código mesmo, que dá alguma saída no terminal e você procura [...].	Problemas no treino do modelo identificado através da análise de log.	Atividade [Treino do modelo]
o meu caso é um pouco mais complicado, porque a aprendizagem é não supervisionada, então não tenho respostas na hora [...]. Eu acho que essa área de aprendizagem não supervisionada está bem atrás da supervisionada.	Overfitting não explorado devido a característica da aprendizagem não supervisionada.	Desafio [Avaliação do modelo]
então, eu uso: Python, Spider, Jupyter, e as bibliotecas como numpy, pandas e o Sklern.	Uso de python, jupyter e frameworks para criação de modelos.	Atividade [Treino do modelo]
como eu te disse, a maioria dos problemas são de implementação, então corrijo buscando documentação, StackOverflow e vou corrigindo.	Problemas de implementação solucionados em foruns.	Atividade [Treino do modelo]
	Problemas de implementação solucionados através da documentação.	Atividade [Treino do modelo]
logo quando eu entrei no projeto, tinha uma planilha gigante, que já tinha os dados no formato que eu usava. Só que aí, quando foi mais tarde com a necessidade de implantar, eu tive que consumir os dados direto do banco, então não tinha eles mais nesse formato certinho.	Dados foram recebidos processados	Limitação [Processamento de dados]
	A implantação do modelo faz necessário modificar o formato de dados recebidos pelo modelo.	Limitação [Avaliação do modelo]
no meu caso [...] teve muita coisa que eu tive que adaptar, tipo, tem o algoritmo do K-Means, só que o K-Means não se dá muito bem com dados faltantes, então eu tive que adaptar para fazer uma variação, só que aí eu aproveitei o sklearn [...].	Modificação de algoritmos de frameworks para novos tipos de problema.	Atividade [Treino do modelo]