

Trecho	Código	Categoria
basicamente seria isso mesmo, os requisitos do modelo seriam o processo que a gente iria automatizar utilizando inteligência artificial.	Requisitos do Modelo como definição do processo a ser automatizado.	Atividade [Requisitos do modelo]
O processamento de dados seria o principal	Processamento de Dados como etapa principal.	Limitação [Processamento de dados]
Os atributos a gente tem que avaliar	Atributos devem ser analisados.	Atividade [Engenharia de atributos]
porque pode ser que alguns dependendo do caso não influenciam, ou [...] podem ser muito pesados e a gente pode reduzir	Atributos podem estar não otimizados.	Limitação [Engenharia de atributos]
então valida e vê se tem que mudar alguma coisa	Incerteza no resultado de treino	Desafio [Treino do modelo]
	Avaliação do modelo indica a correteude do treino.	Limitação [Avaliação do modelo]
é como se a gente (refere-se a sua equipe de trabalho) já tivesse o objetivo, tem que fazer um classificador de...	Requisitos do Modelo como definição do modelo a ser criado.	Limitação [Requisitos do modelo]
Nós trabalhamos em módulos isolados [...] esses módulos serão usados em sistema maior que tem todo processo de negócio a ser seguido [...] então fica meio que independente.	Modelos de machine learning como módulos em um MLES.	
a gente (refere-se a empresa) tem um sistema a desenvolver, uma ferramenta, no final a gente entrega a solução inteira.	Empresa desenvolve MLES.	
eu estou envolvido em todas as etapas	Único desenvolvedor em todas as etapas.	Ator [Geral]
mas às vezes dependendo do projeto tem que ter o especialista nos dados (refere-se a um especialista no domínio da aplicação, não um especialista em data science), um pessoa que [...] consiga interpretar os dados, que tenha conhecimento sobre aquela área de negócio [...].	Presença de um especialista de domínio.	Ator [Geral]
Eu trabalho com imagem, às vezes eu preciso classificar coisas na imagem e eu não sei o que são, então eu dependo de um especialista que vai anotar as imagens pra mim [...].	Desenvolvedor incapaz de reconhecer seus dados	Desafio [Processamento de dados]
	Especialista de domínio anota os dados	Atividade [Processamento de dados]
Depende, em casos que você tem literatura você já tem um alvo a atingir, você tem que chegar naquela medida de acurácia,	Literatura define as metas de avaliação.	Limitação [Avaliação do modelo]
mas dependendo do caso, se você estiver trabalhando em um negócio novo, qualquer coisa seria melhor, acima de 50% ou 60%.	Novas soluções não precisam de ótimos resultados de avaliação.	Limitação [Avaliação do modelo]
não chegamos assim a fundo nesse ponto [...], é só mais a acurácia mesmo [...]. A depender, em na detecção de imagens também uso o recall.	Uso de métricas clássicas.	Método [Avaliação do modelo]
	Métricas clássicas utilizadas na avaliação do modelo.	Atividade [Avaliação do modelo]
o workflow se repete, pode se repetir sim [...]. Porque você não tem certeza que aquilo é o caso ótimo.	Ciclos de desenvolvimento existem devido a incerteza de correteude.	Desafio [Geral]
depende da especificidade de cada projeto, mas aparentemente o processamento de dados.	O projeto influencia na etapa que mais se repete.	Limitação [Geral]
	Processamento de dados como etapa que mais se repete.	Desafio [Processamento de dados]
Por exemplo, na detecção de imagem nós estávamos trabalhando (refere-se a uma atividade anterior) com um dataset que não era tão grande, então (a princípio) fizemos um treino com o que tinha. Na avaliação ficou mais ou menos, aí eu voltei para o processamento de dados, gerei mais imagens: girei, inseri ruído [...].	Avaliação da indícios da correteude do modelo.	Limitação [Avaliação do modelo]
	Enriquecimento de dados.	Método [Processamento de dados]
	Dados em quantidade ou qualidade inferior são enriquecidos.	Atividade [Processamento de dados]
	Processamento de dados como etapa mais problemática	Desafio [Processamento de dados]
Processamento de dados, sem dúvidas, porque tudo depende disso, se os dados estiverem errados, o modelo vai sair errado [...].	Dados errados implicam em modelo errado.	Limitação [Processamento de dados]
	Quantidade dos dados como aspecto importante	Limitação [Processamento de dados]
Não só a quantidade dos dados, mas também a qualidade, pra balancear [...]. Então é o principal a definir para fazer qualquer coisa.	Qualidade dos dados como aspecto importante.	Limitação [Processamento de dados]
	Dificuldade em concretizar o dataset	Desafio [Processamento de dados]
esse é o principal problema que eu vejo em Machine Learning, montar o dataset [...], ter certeza que aquilo tá certo (refere-se aos dados).	Dificuldade em garantir a correteude do dataset.	Desafio [Processamento de dados]
	O projeto influencia nos problemas a serem enfrentados	Limitação [Desenvolvimento]
depende de cada caso, no meu caso em específico é a dependência que tenho do especialista (refere-se ao especialista no domínio da aplicação)	Desenvolvedor incapaz de reconhecer seus dados	Desafio [Processamento de dados]

no meu caso há todo um processo de aquisição das imagens, então eu identifico visualmente	Identificação visual de problemas em imagens	Atividade [Processamento de dados]
em alguns projetos (refere-se aos de não detecção de imagens) são utilizadas ferramentas para gerar gráficos dos dados como histogramas e boxplots, geralmente com o auxílio do especialista.	Gráficos utilizados para identificação dos problemas	Atividade [Processamento de dados]
	Uso de gráficos para a visualização de dados: boxplot, histogramas.	Método [Processamento de dados]
	Especialista de domínio para a identificação de problemas.	Atividade [Processamento de dados]
	Presença de um especialista de domínio.	Ator [Processamento de dados]
eu vejo sempre na literatura (se alguma ferramenta é indicada), mas nenhuma em específico.	Literatura define novas ferramentas a serem incorporadas	Limitação [Geral]
sabendo que os dados estão corretos, se o especialista me diz que naquela região X tem aquilo [...]. Eu só tenho que ter amostras e testar arquiteturas.	Especialista de domínio anota os dados	Atividade [Processamento de dados]
	Desenvolvedor confia no especialista de domínio.	Limitação [Processamento de dados]
isso depende de análise específica do modelo que você está utilizando, mas é mais empírico sim, analisar a saída camada a camada e testar [...].	Definição empírica de hiperparâmetros para os modelos	Atividade [Treino do modelo]
a gente já tem bibliotecas que tem a arquitetura implementada, até porque as arquiteturas para redes convolucionais são bem grandes, então fazer do zero ...	Uso de frameworks para criação de modelos.	Método [Treino do modelo]
no meu caso, eu não sei. Porque os desafios que nós temos hoje [...] são de detecção de objetos do mundo real [...]. Eu estou em um outro contexto que não tem muito a ver com isso, então eu utilizo pesos pré-treinados. [...] Pode ser que seja melhor fazer do zero ou não.	Incerteza sobre a correteude dos modelos.	Desafio [Treino do modelo]