Soluções de Aprendizagem de Máquina

Introdução

Agenda

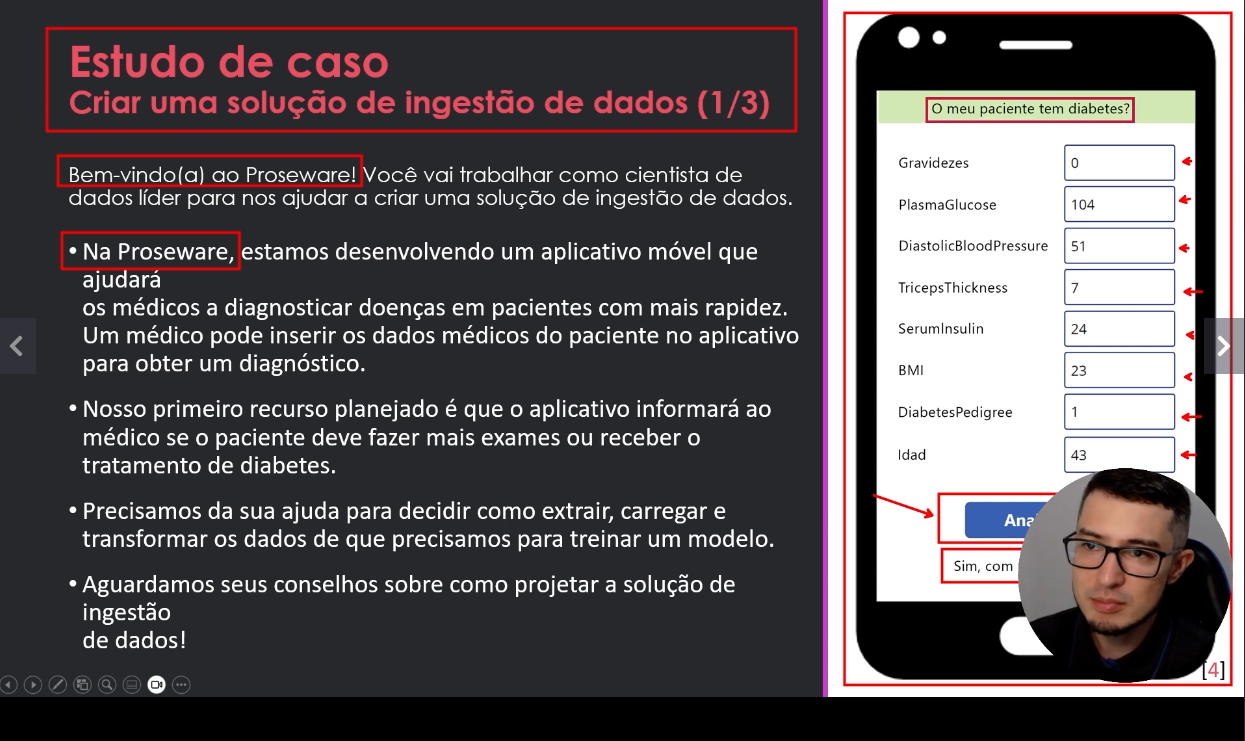
* Criar uma solução de ingestão de dados para projetos de aprendizado de máquina
* Criar uma solução de treinamento de modelo de machine learning
* Criar uma solução de implantação de modelo
* Criar uma MLOPs (Solução de operações de aprendizado de máquina)

Estudo de Caso: Criar uma solução de ingestão de dados

Caso de uso

Criar uma solução de ingestão de dados

Bem-vindo ao Proseware! Você vai trabalhar como cientista de dados líder para nos ajudar a criar uma solução de ingestão de dados.

* Na Proseware, estamos desenvolvendo um aplicativo móvel que ajudará os médicos a diagnosticar doenças em pacientes com mais rapidez para obter um diagnóstico.
* Nosso primeiro recurso planejado é que o aplicativo informará ao médico se o paciente deve fazer mais exames ou receber o tratamento de diabetes.
* Precisamos da sua ajuda para decidir como extrair, carregar e transformar os dados de que precisamos para treinar um modelo.
* Aguardamos seus conselhos sobre como projetar a solução de ingestão de dados!  
  

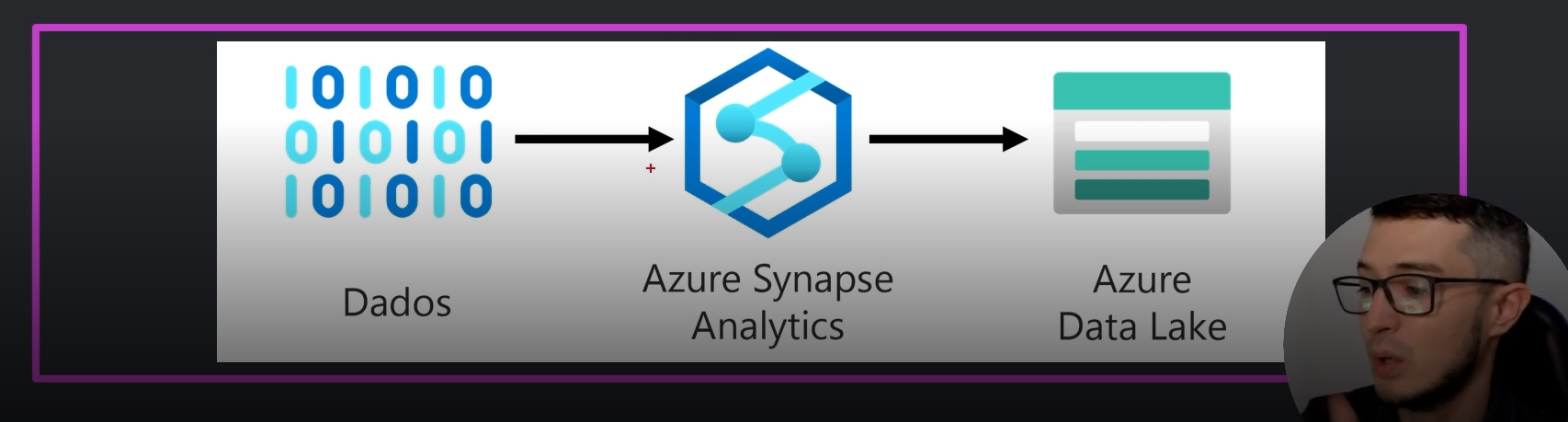
Considere os requisitos:

* Considerar o tipo de dados atual: já coletamos dados correlacionados com diabetes, com o número de gestações, idade e imc (índice de massa corporal) em um banco de dados de pacientes.
* Considerar o tipo de dados desejado: nossos cientistas de dados estão acostumados a trabalhar com python e querem os dados como arquivos csv.
* Considerar o acesso aos dados: querermos que nosso design esteja preparado o futuro e pronto para escala. Queremos extrair os dados confidenciais de privacidade do banco de dados do paciente e armazená-los em uma solução de armazenamento de dados Azure.

Qual solução de armazenamento você recomendaria para armazenar os dados? Data Lake gen 2

Qual ferramenta você recomendaria que usássemos pra mover os dados

? Azure synapse analytics



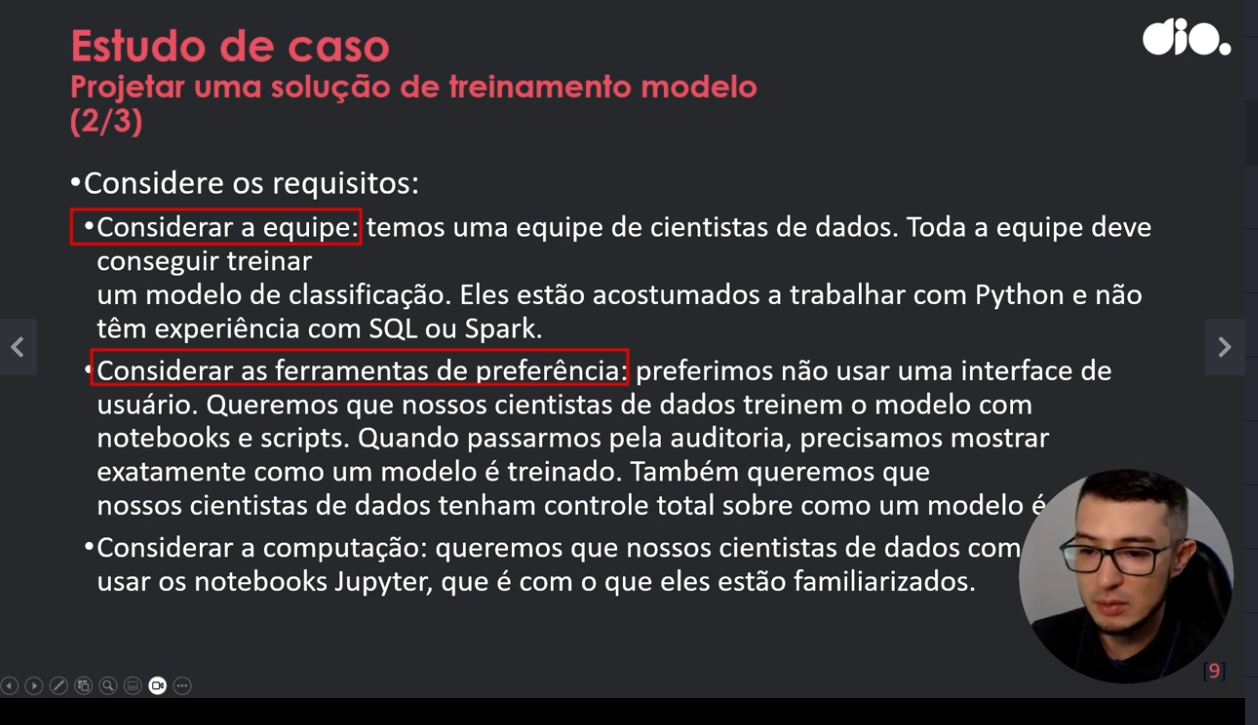
Estudo de caso: Como projeto uma solução de treinamento modelo de machine learning

Bem-vindo ao Proseware! Você vai trabalhar como cientista de dados líder para nos ajudar a criar uma solução de ingestão de dados.

* Na Proseware, estamos desenvolvendo um aplicativo móvel que ajudará os médicos a diagnosticar doenças em pacientes com mais rapidez para obter um diagnóstico.
* Nosso primeiro recurso planejado é que o aplicativo informará ao médico se o paciente deve fazer mais exames ou receber o tratamento de diabetes.
* Precisamos de sua ajuda para decidir como treinar um modelo que possa detectar o diabetes.
* Aguardamos seus conselhos sobre como projetar a solução de treinamento de dados!

Considere os requisitos:

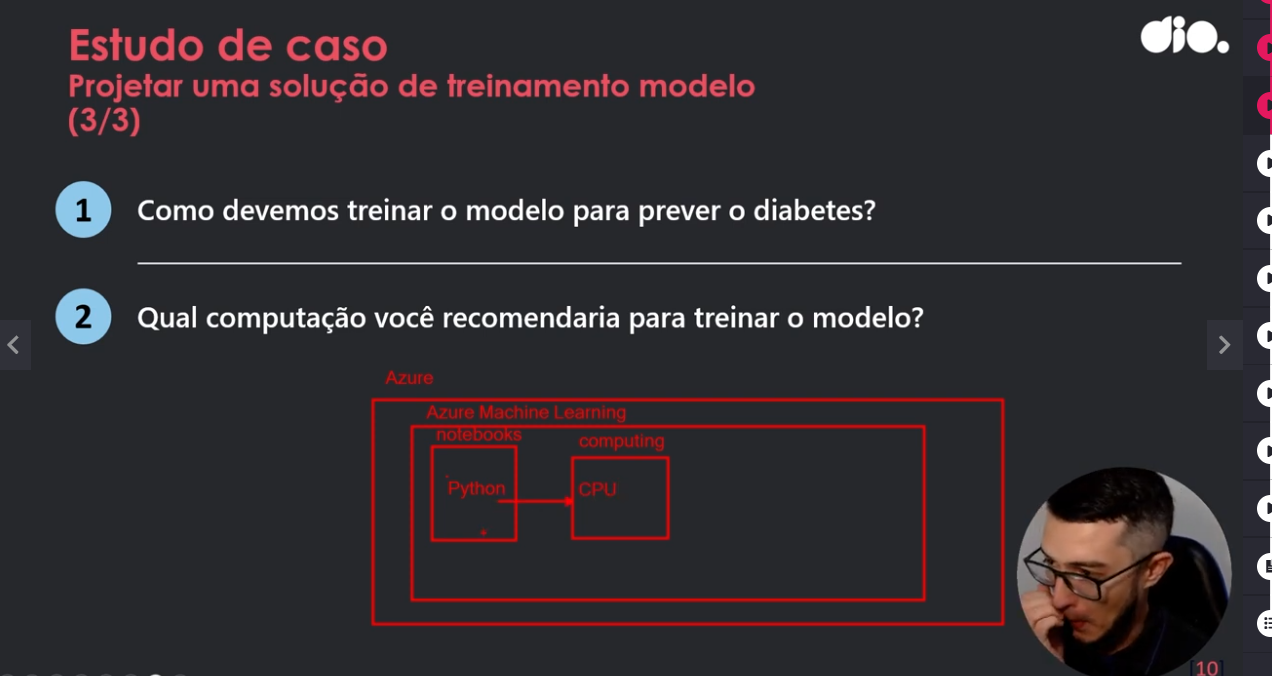
* Considerar a equipe: Temos uma equipe de cientista de dados. Toda a equipe deve conseguir treinar um modelo de classificação. Eles estão acostumados a trabalhar com Python e não tem experiência com SQL ou Spark.
* Considerar as ferramentas de preferência: Preferimos não usar uma interface de usuário. Queremos que nossos cientistas de dados treinem o modelo com notebooks e scripts. Quando passarmos pela auditoria, precisamos mostrar exatamente como um modelo é treinado. Também queremos que nossos cientistas de dados tenham controle total sobre como um modelo é
* Considerar a computação: Queremos que nossos cientistas de dados como usar os notebooks jupyter, que é com o que eles estão familiarizados.



Como devemos treinar o modelo para prever o diabetes?

Azure Machining Learning

Qual computação você recomendaria para reinar o modelo?  
Computing CPU



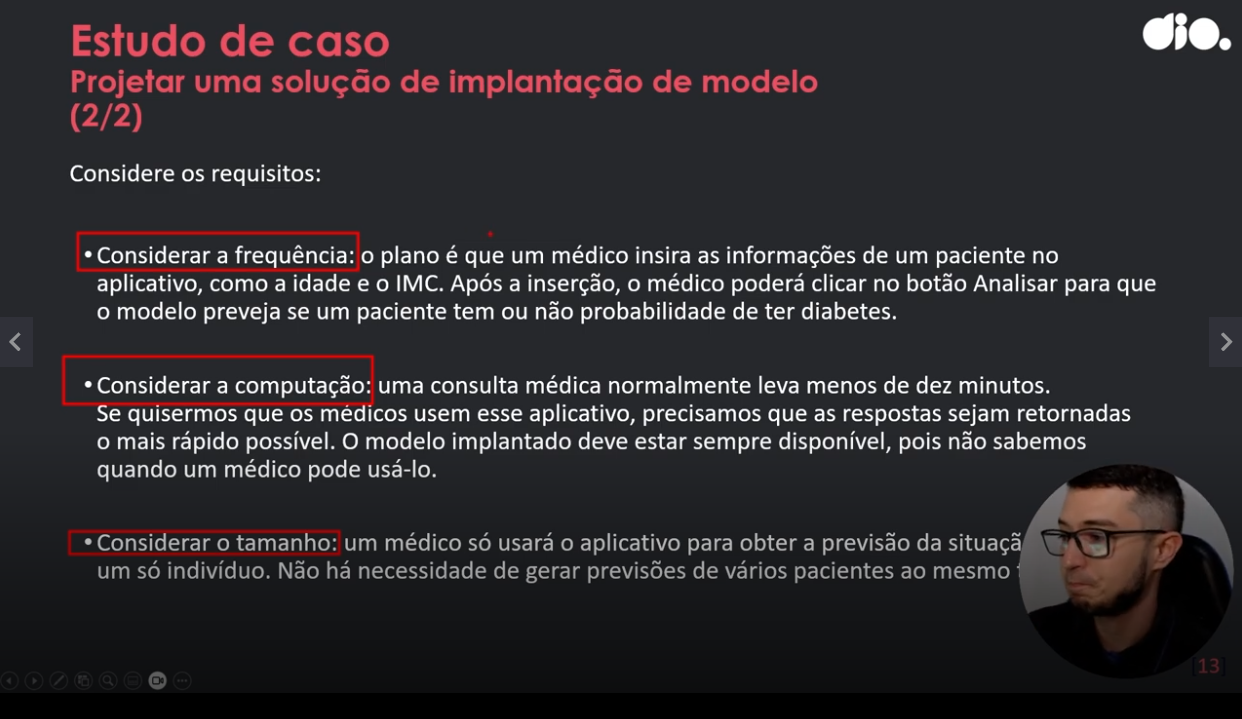
Estudo de caso: Projetar uma solução de implantação de modelo

Bem-vindo ao Proseware! Você vai trabalhar como cientista de dados líder para nos ajudar a criar uma solução de ingestão de dados.

* Na Proseware, estamos desenvolvendo um aplicativo móvel que ajudará os médicos a diagnosticar doenças em pacientes com mais rapidez para obter um diagnóstico.
* Nosso primeiro recurso planejado é que o aplicativo informará ao médico se o paciente deve fazer mais exames ou receber o tratamento de diabetes.
* Precisamos da sua ajuda para decidir como implantar o modelo para integrá-lo ao aplicativo móvel
* Aguardamos seus conselhos sobre como projetar a solução de implantação do modelo.

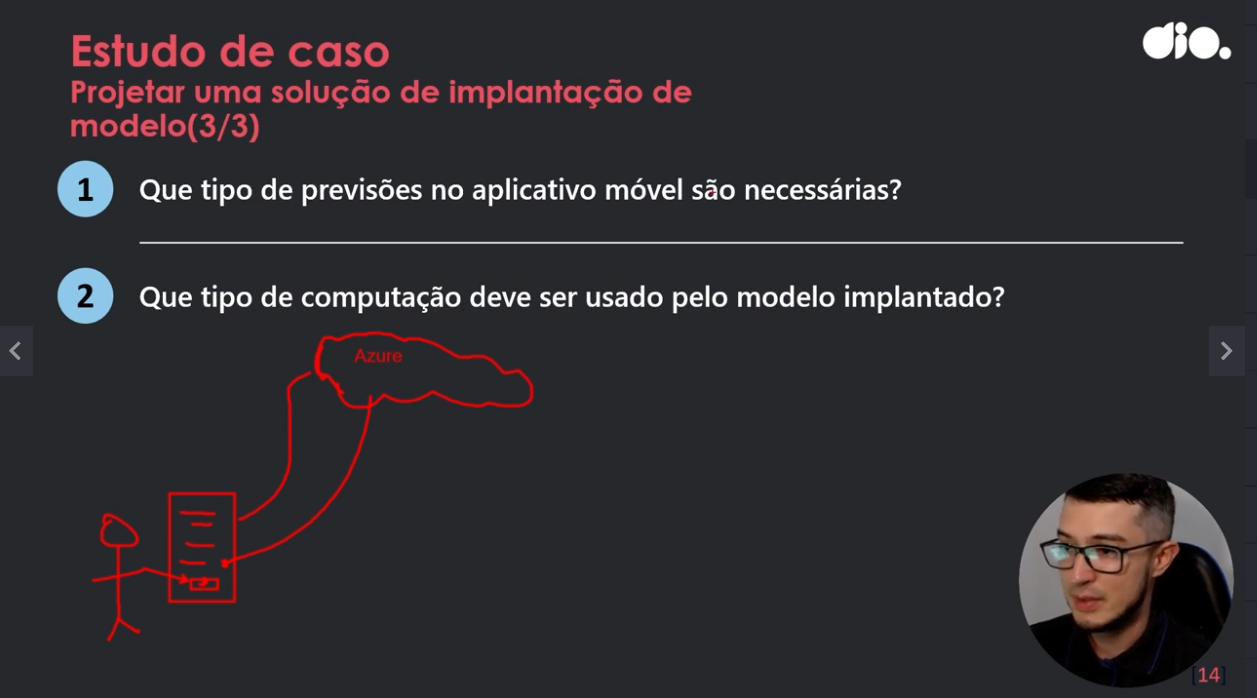
Considere os requisitos:

* Considerar a frequência: o plano é que um médico insira as informações de um paciente no aplicativo, com a idade e o IMC. Após a inserção, o médico poderá clicar no botão Analisar para que o modelo preveja se um paciente tem ou não probabilidade de ter diabetes.
* Considerar a computação: Uma consulta médica normalmente leva menos de dez minutos. Se quisermos que os médicos usem esse aplicativo, precisamos que as respostas sejam retornadas o mais rápido possível. O modelo implantado deve estar sempre disponível, pois não sabemos quando um médico pode usá-lo.
* Considerar o tamanho: Um médico só usará o aplicativo para obter a previsão da situação um só indivíduo. Não há necessidade de gerar previsões de vários pacientes ao mesmo tempo.



Que tipo de previsões no aplicativo móvel são necessárias?

Em tempo real.

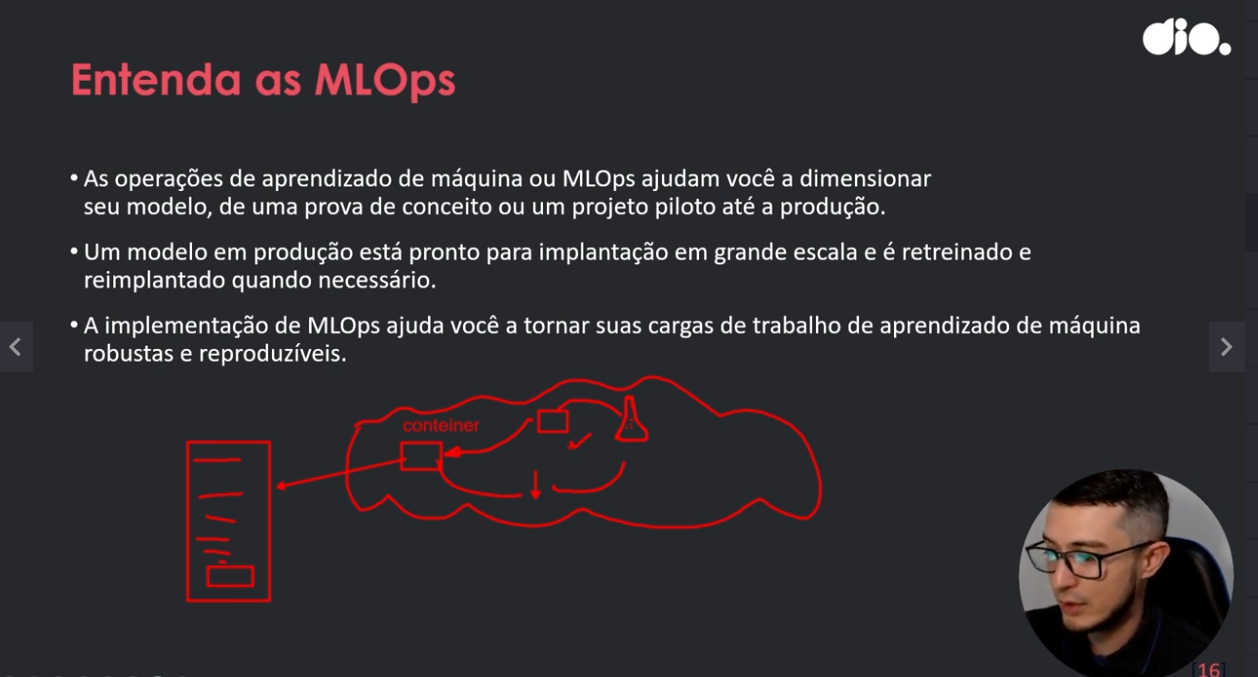


Que tipo de computação deve ser usado pelo modelo implantado?

Container.

Criar uma MLOps (Solução de operações de aprendizado de máquina)

Entenda as MLOps

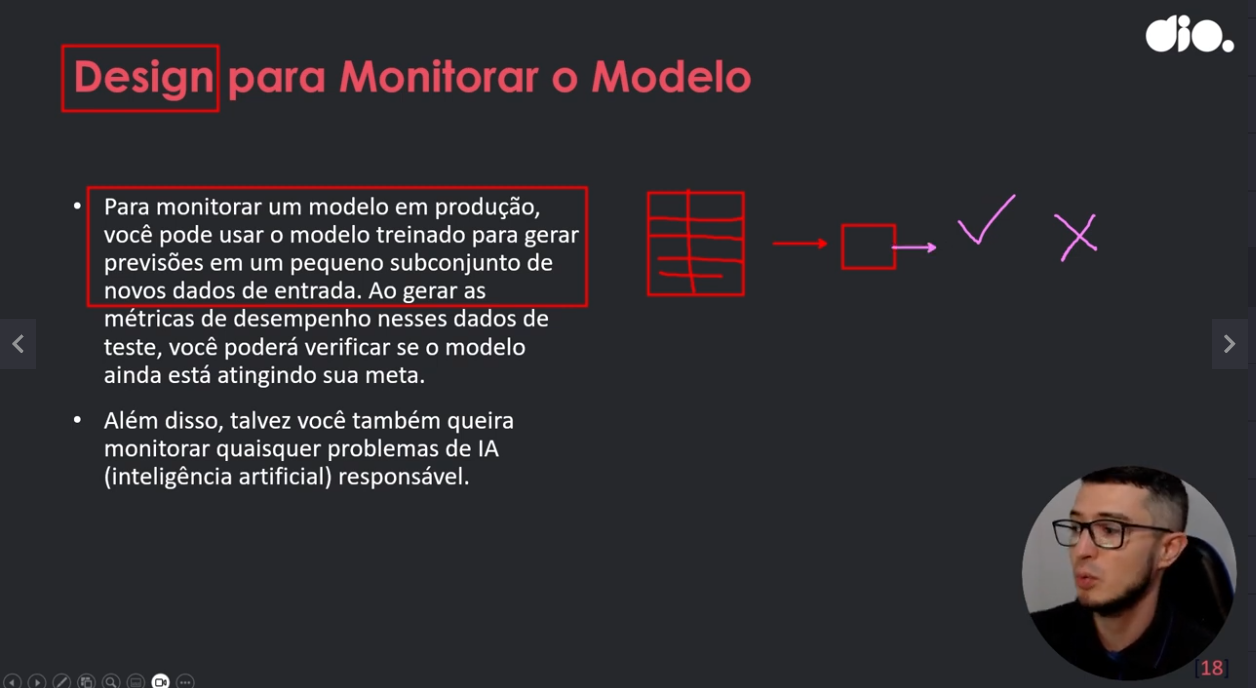
* As operações de aprendizado de máquina ou MLOps ajudam você a dimensionar seu modelo, de uma prova de conceito ou um projeto piloto até a produção.
* Um modelo em produção está pronto implantação em grande escala e é retreinado e reimplantado quando necessário.
* A implementação de MLOps ajuda você a tornar suas cargas de trabalho de aprendizado de máquina robustas e reproduzíveis. 

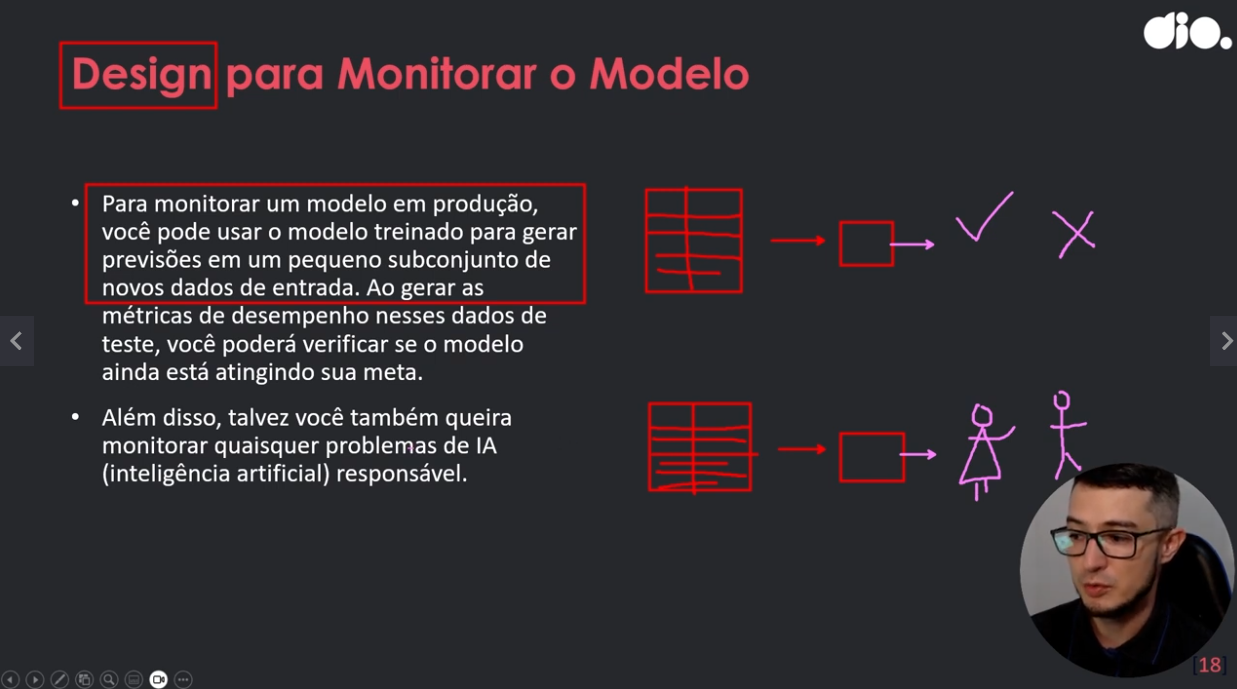
No MLOps (e DevOps), um ambiente constitui uma coleção de recursos.

* Esses recursos são usados para implantar um aplicativo ou, em projetos de aprendizado de máquina, implantar um modelo. Em um projeto MLOps, eles se referem aos recursos do Azure necessários para uma fase do projeto.



Design para monitorar o modelo

* Para monitorar um modelo em produção, você pode usar o modelo treinado para gerar previsões em um pequeno subconjunto de novos dados de entrada. Ao gerar as métricas de desempenho nesses dados de teste, você poderá verificar se o modelo ainda está atingindo sua meta.
* Além disso, talvez você também queira monitorar quaisquer problemas de IA responspavel



Design para Monitorar os Dados

* Normalmente, você treina um modelo de machine learning usando um conjunto de dados de histórico que representa os novos dados que seu modelo receberá quando implantado. No entanto, ao longo do tempo pode haver tendências que alteram o perfil dos dados, tornando o modelo menos preciso.
* Essas alterações nos perfis de dados entre os dados de treinamento e atuais é conhecida como descompasso de dados e pode ser um problema significativo para modelos de previsão usados na produção.



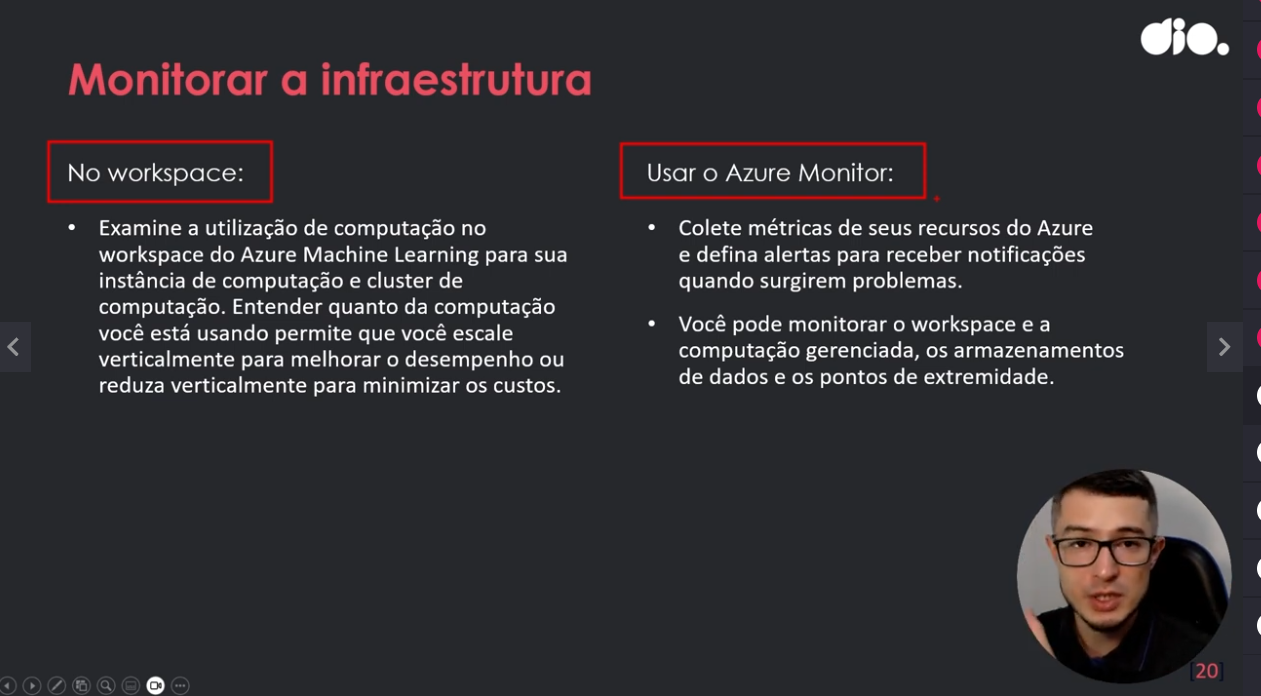
Monitorar a infraestrutura.

No workspace:

* Examine a utilização de computação no workspace do Azure Machine Learning para sua instância de computação e cluster de computação. Entender quanto da computação você está usando permite que você escale verticalmente para melhorar o desempenho ou reduza verticalmente para minimizar os custos.

Usar o Azure Monitor:

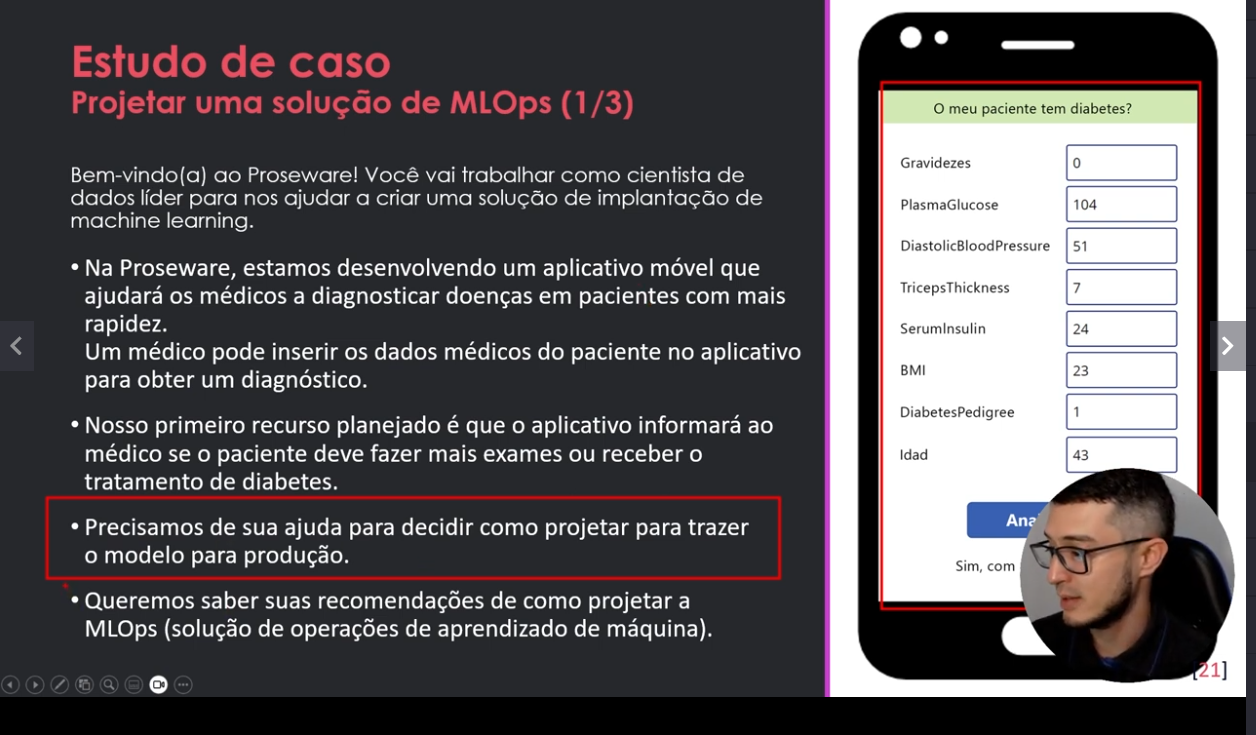
* Colete métricas de seus recursos do Azure e defina alertas para receber notificações quando sugerem problemas.
* Você pode monitorar o workspace e a computação gerenciada, os armazenamentos de dados e os pontos de extremidade.



Estudo de caso: Projetar uma solução de MLOps

Bem-vindo ao Proseware! Você vai trabalhar como cientista de dados líder para nos ajudar a criar uma solução de ingestão de dados.

* Na Proseware, estamos desenvolvendo um aplicativo móvel que ajudará os médicos a diagnosticar doenças em pacientes com mais rapidez para obter um diagnóstico.
* Nosso primeiro recurso planejado é que o aplicativo informará ao médico se o paciente deve fazer mais exames ou receber o tratamento de diabetes.
* Precisamos da sua ajuda para decidir como projetar para trazer o modelo para produção.
* Queremos saber suas recomendações de como projetar a MLOps (Solução de operações de aprendizado de máquina).



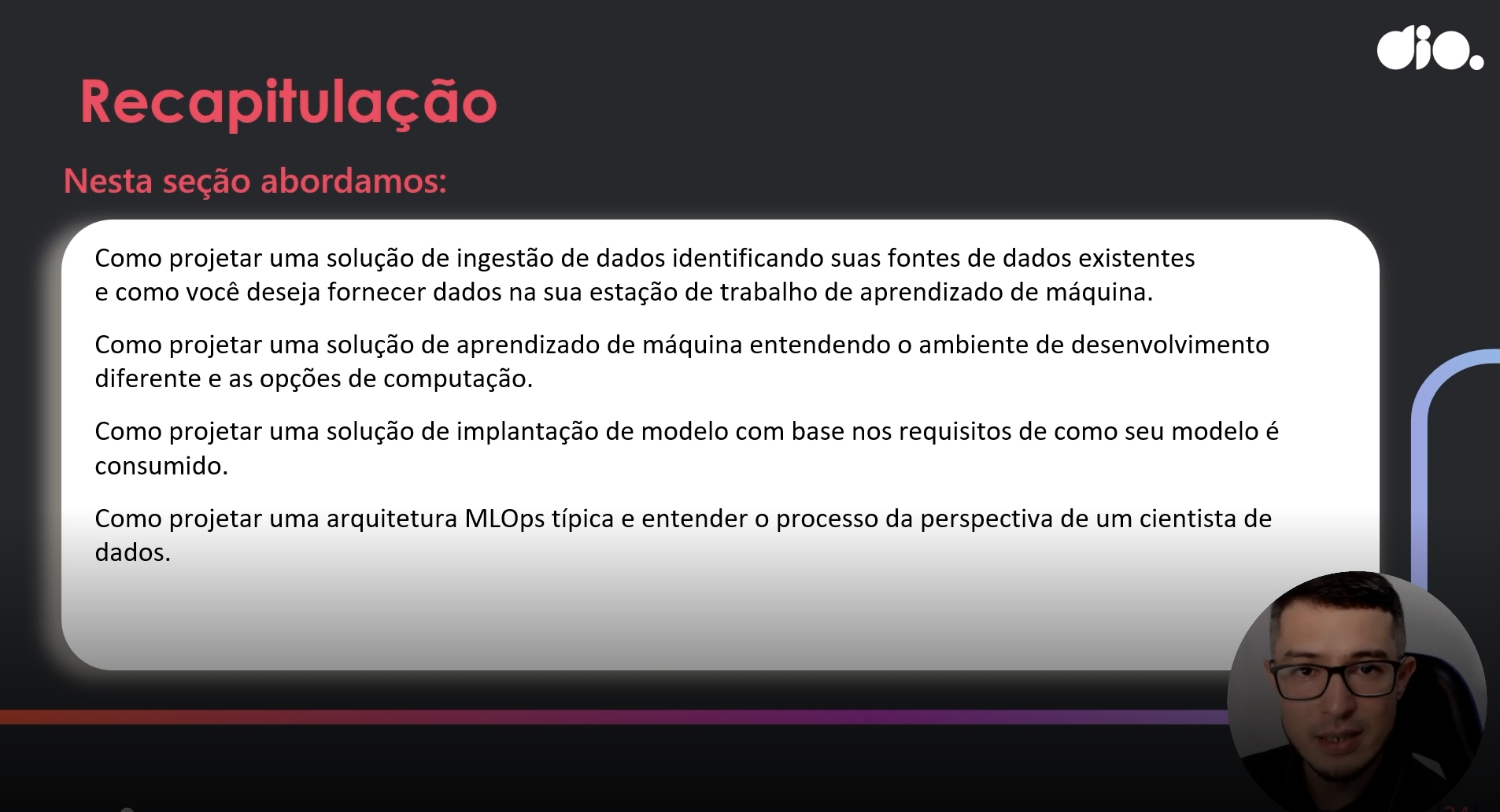
Considere os requisitos:

* Considere os ambientes: atualmente, estamos trabalhando em uma pequena equipe e você é a única pessoa cientista de dados envolvida. Queremos ver se esse projeto é bem-sucedido antes de realmente escalar verticalmente e envolver uma grande equipe.
* Considere o modelo: Como o modelo é usado para ajudar os médicos, a precisão é importante para nós. O modelo só deve estar um uso quando soubermos que ele está funcionando conforme o esperado.
* Considere os dados: Estamos começando pequeno e usaremos principalmente o modelo implantado para testar nosso aplicativo. os dados nos quais o modelo implantado gera previsões não devem ser usados para treinar novamente o modelo, pois ele pode ser tendencioso

Quantos workspaces do Azure Machine Learning a equipe deve criar?

Somente 1

Quando devemos treinar novamente o modelo? Nesse caso precisa verificar a performance, ou se está saindo dos parâmetros de acurácia estabelecidos



**Materiais de Apoio**

Os materiais complementares e de apoio que oferecemos têm como objetivo fornecer informações para facilitar e enriquecer a sua jornada de aprendizado no curso "**Projetar uma Solução de Aprendizado de Máquina**". Aqui você encontrará links úteis, como slides, repositórios e páginas oficiais, além de dicas sobre como se destacar na DIO e no mercado de trabalho 😉

**Recursos Adicionais**

Durante este conteúdo, compreendemos os fundamentos da engenharia de prompts. Para ajudá-lo a aprofundar o conhecimento, disponibilizamos a seguir o material complementar contendo os conteúdos e links apresentados no curso:

* **Slide**: [Projetar uma Solução de Aprendizado de Máquina.pptx](https://hermes.dio.me/files/assets/53a57c15-c099-4fa4-8df6-c612e14b86d0.pptx)

**Dicas e Links Úteis**

Para se desenvolver ainda mais e se destacar na DIO e no mercado de trabalho, sugerimos os seguintes recursos:

* **Artigos e Fórum da DIO**: Compartilhe seus conhecimentos e dúvidas através dos [artigos](https://web.dio.me/articles) (visíveis globalmente na plataforma da DIO) e nos fóruns específicos para cada experiência educacional, como nossos Bootcamps.
* **Rooms**: Participe do *Rooms*, uma ferramenta de bate-papo em tempo real onde você pode interagir com outros participantes dos nossos Bootcamps, compartilhando dúvidas, dicas e snippets de código.
* **Exploração na Web**: Utilize motores de busca para aprofundar seu conhecimento sobre temas específicos. Páginas como o [StackOverflow](https://stackoverflow.com/) são recursos valiosos para encontrar soluções e expandir seu entendimento.

Com esses materiais complementares, você estará bem equipado para explorar todo o potencial e se destacar em suas iniciativas. Continue aproveitando as oportunidades de aprendizado, e não hesite em buscar mais conhecimento e compartilhar suas descobertas com a comunidade!

