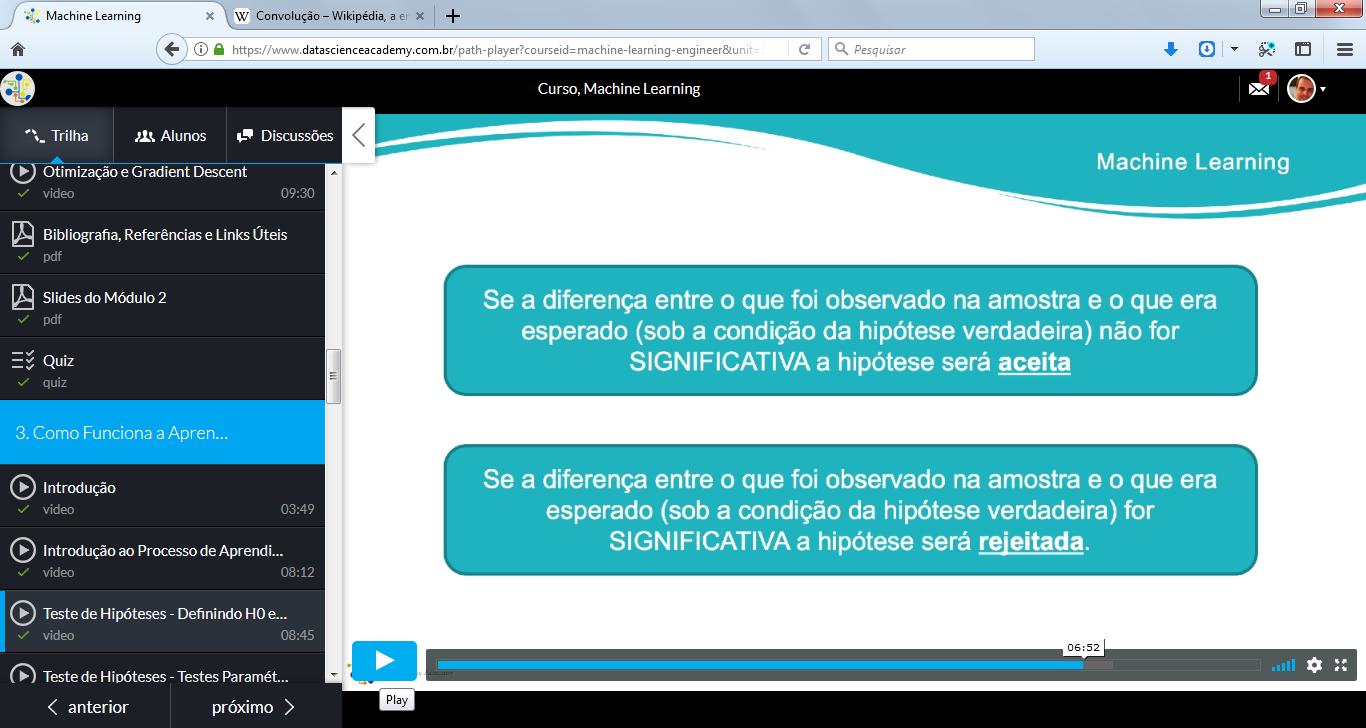
**Teste de Hipóteses**

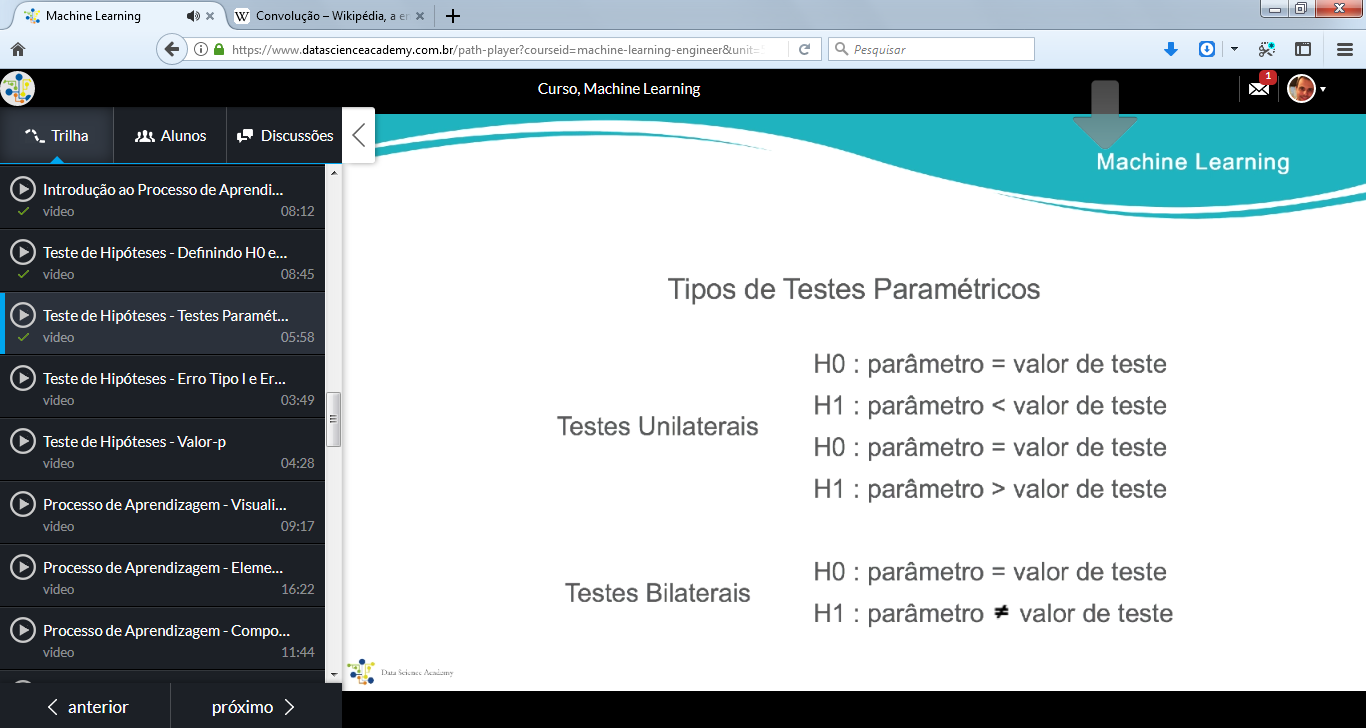
**H0 = Hipótese Nula** - Aceita como verdadeira até que se prove o contrário

**H1 = Hipótese Alternativa**

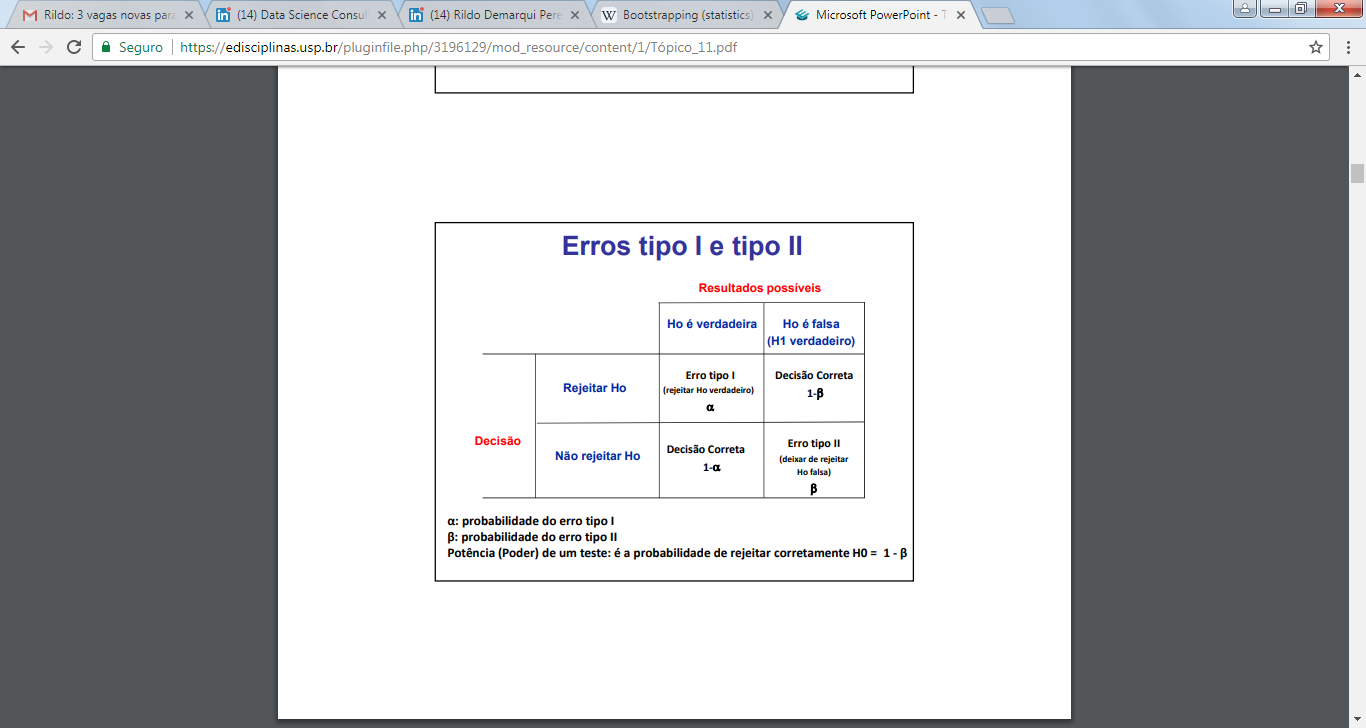
Uma hipótese estatística é uma suposição ou afirmação que pode ou não ser verdadeira, relativa a uma ou mais populações. A veracidade ou falsidade de uma hipótese estatística nunca é conhecida com certeza, a menos que, se examine toda a população, o que é impraticável na maior parte das situações. Desta forma, toma-se uma amostra aleatória da população de interesse e com base nesta amostra é estabelecido se a hipótese é provavelmente verdadeira ou provavelmente falsa.



**Tipos de Testes Paramétricos**



**Tipos de Erros**



**Erro Tipo I -** Rejeitar H0 quando de fato H0 for verdadeiro (alfa)

**Erro Tipo II -** Não rejeitamos H0 quando de fato H0 é falso (beta)

**Nível de Significancia (alfa)**

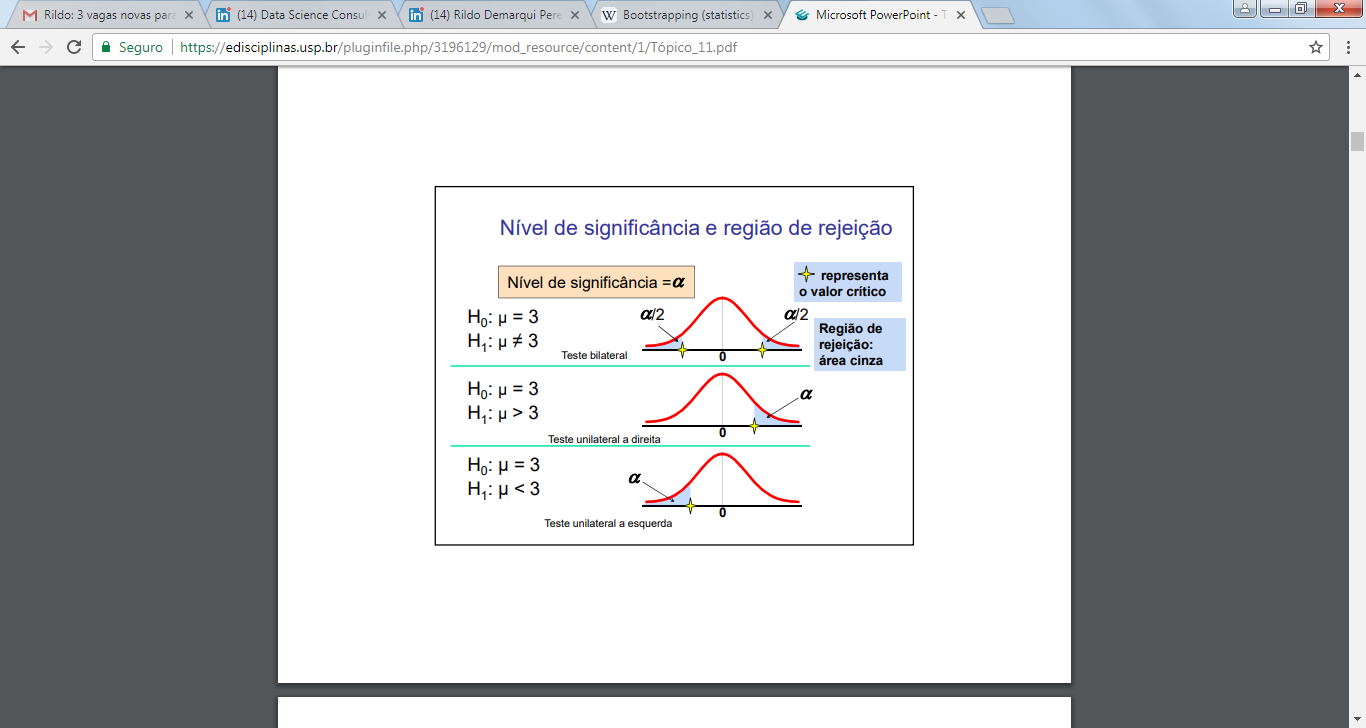
A probabilidade de cometermos um erro tipo I é chamada de nível de significância, denotado pela letra alfa.

Define os valores improváveis da estatística amostral se a hipótese nula é verdadeira, ou seja, a região de rejeição da distribuição amostral.

Valores típicos para alfa: 0.01, 0.05, or 0.10 (é uma probabilidade)

Deve ser selecionada antes do teste.

Fornece os valores críticos do teste.



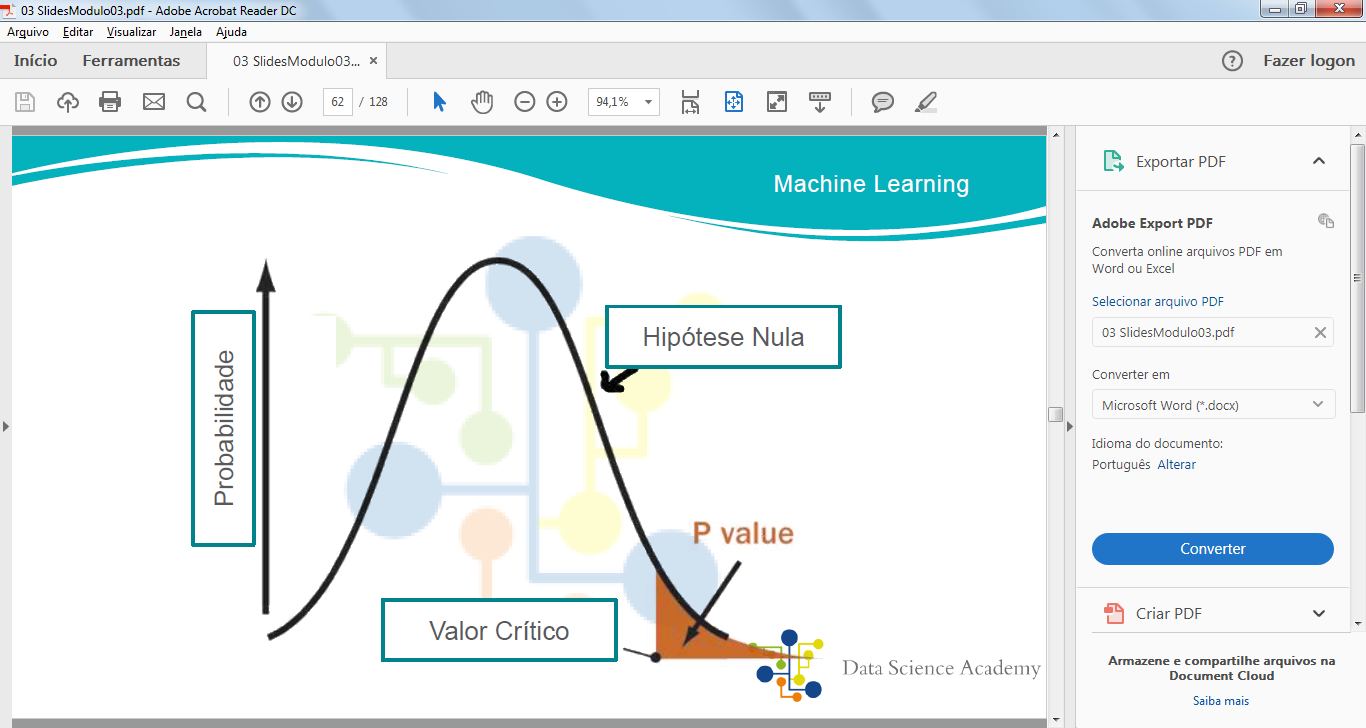
**Valor P (p-value)**

Um p-value é uma medida de quanta evidência você tem contra a hipótese nula. Quanto menor o p-value, mais evidência você tem. Deve-se combinar o p-value com o nível de significância para tomar decisão sobre um dado teste de hipótese. Em tal caso, se o p-value for menor que algum corte (usualmente 0,05, algumas vezes um pouco mais como 0,1 ou um pouco menos como 0,01) então você rejeita a hipótese nula.

Com base nos dados da amostra o teste dermina se a hipótese nula deve ser rejeitada. Utilizamos o valor p para fazer a determinação. Se o valor de P é menor ou igual ao nível de significância, que é o ponto de corte que definimos, então rejeita-se a hiótese nula.

Se p > alfa, não rejeita H0

Se p ≤ alfa, rejeita H0



É preciso muita cautela na interpretação de um p-value, dado que esta medida é bastante influenciada pelo tamanho da amostra. Amostras grandes tendem a produzir p-values pequenos, ainda que o efeito observado não tenha grande importância prática, enquanto amostras pequenas tendem a produzir p-values grandes, ainda que exista um importante efeito em um ponto de vista prático. Por isso, o uso dos valores-p nas pesquisas médicas tem sido bastante criticado por vários autores.

**Elementos do Processo de Aprendizagem**

**Há um padrão**

**Não á um único modelo matemático que explique esse padrão**

**Dados estão disponíves**

**Componentes do Processo de Aprendizagem**

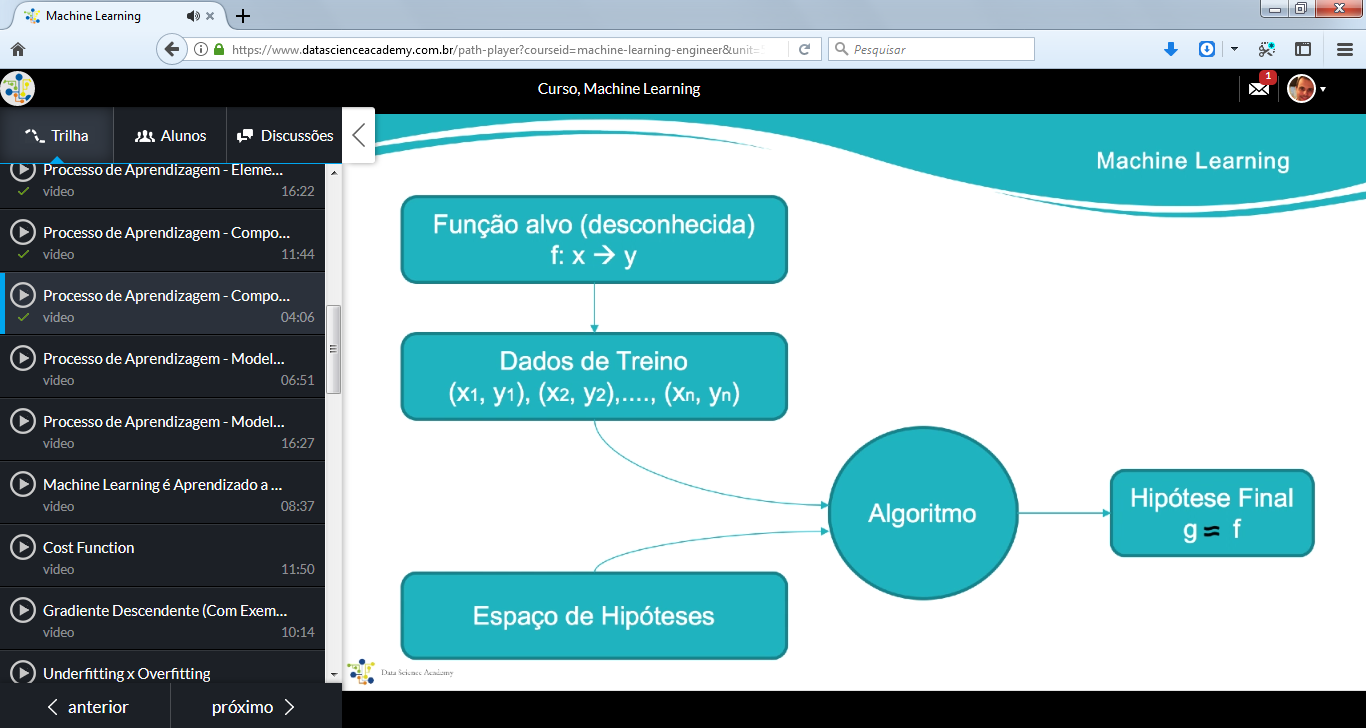
**Input Dados do Cliente**

**Output Decisão -> Crédito: SIM/NÃO**

**Função Alvo Representação do relacionamento**

**Dados Dados históricos**

**Hipótese Fórmula a ser usada**



**Modelo de Aprendizagem = Espaço de Hipótese + Algoritmo de Aprendizagem**

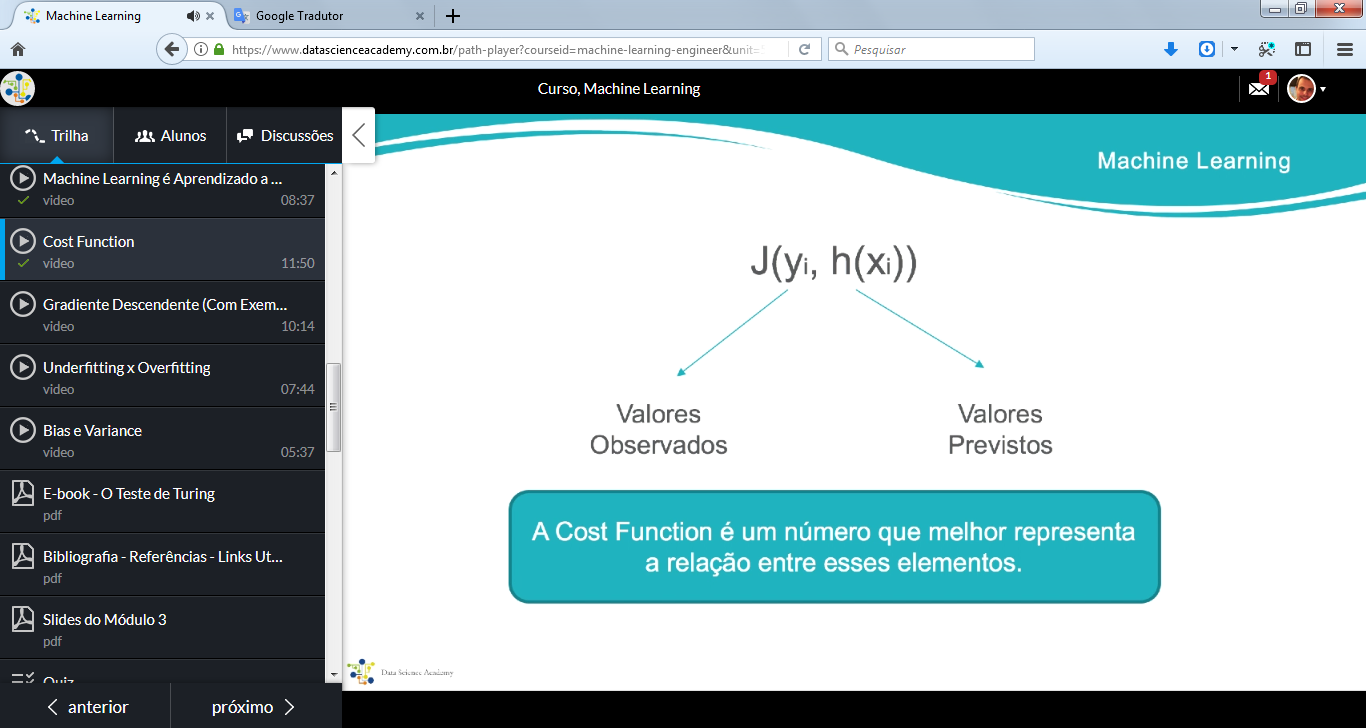
Redes Neurais Back Propagation

Support Vector Machines Programação quadrática

O Espaço de Hipótese contém os recursos com os quais podemos trabalhar. O algoritmo de aprendizagem recebe os dados e navega pelo espaço de hipóteses a fim de encontrar a melhor hipótese que gera o resultado desejado.

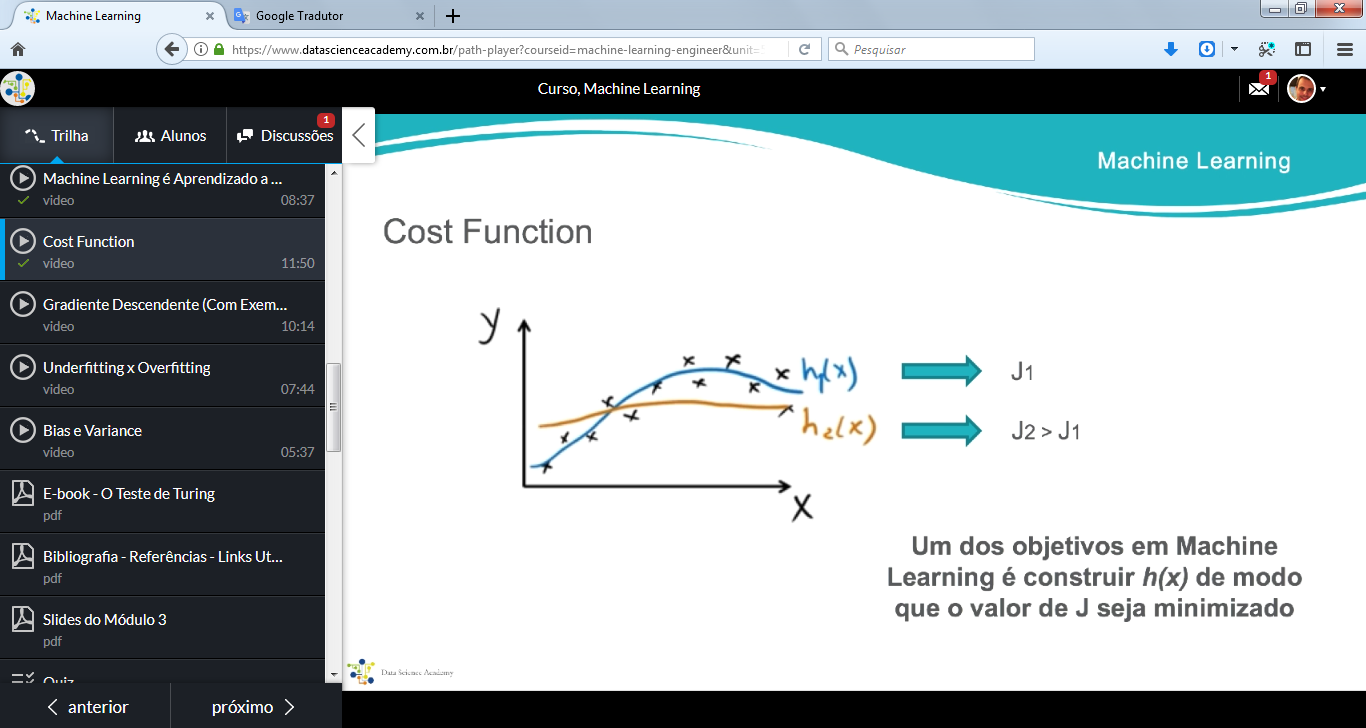
**Cost Function**

Descreve quão bem a resposta na área de resposta (espaço de hipóteses) se encaixa no conjunto de dados que está sendo analisado.



Comparando uma previsão contra o seu valor real, usando uma cost function, determinamos o nivel de erro do algoritmo.

Por ser uma formulação matemática, a cost function expressa o nível de erro em uma forma numérica. A cost function transmite o que é realmente importante e significativo para seus propósitos com o algoritmo de aprendizagem.



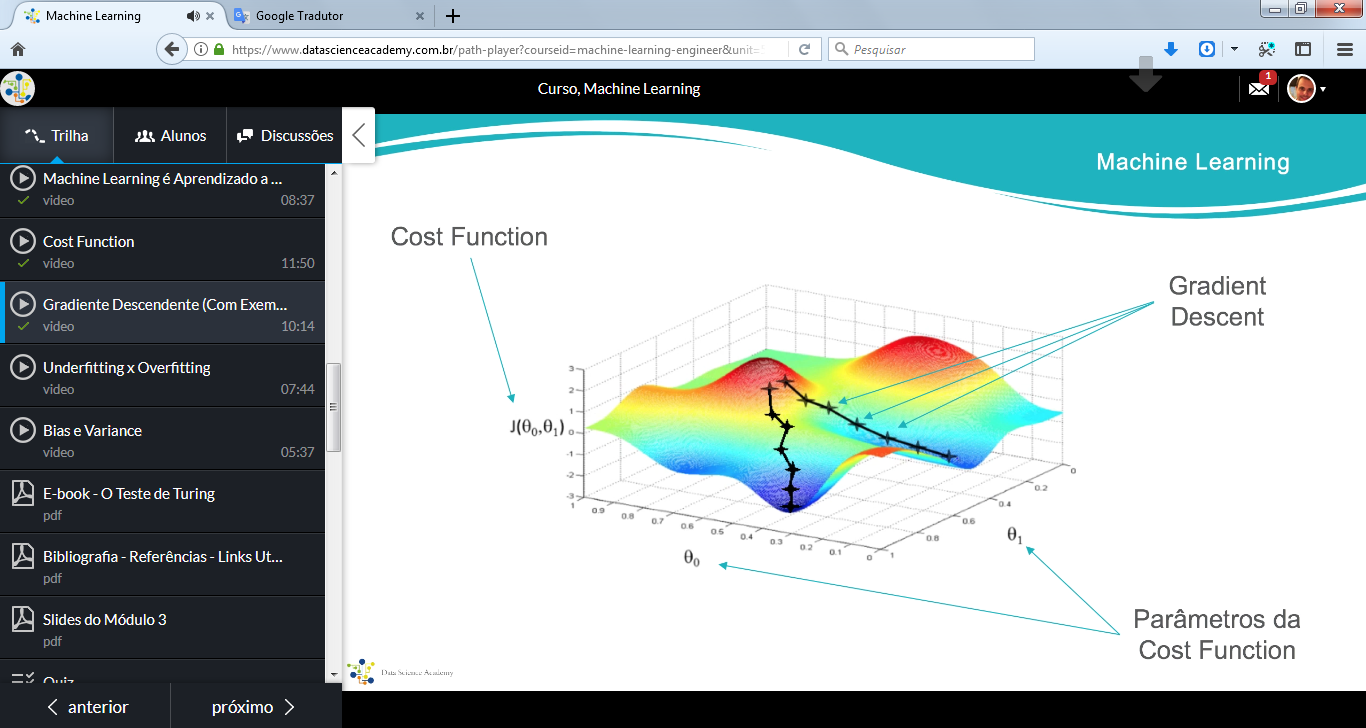
Onde:

h -> Hipóteses

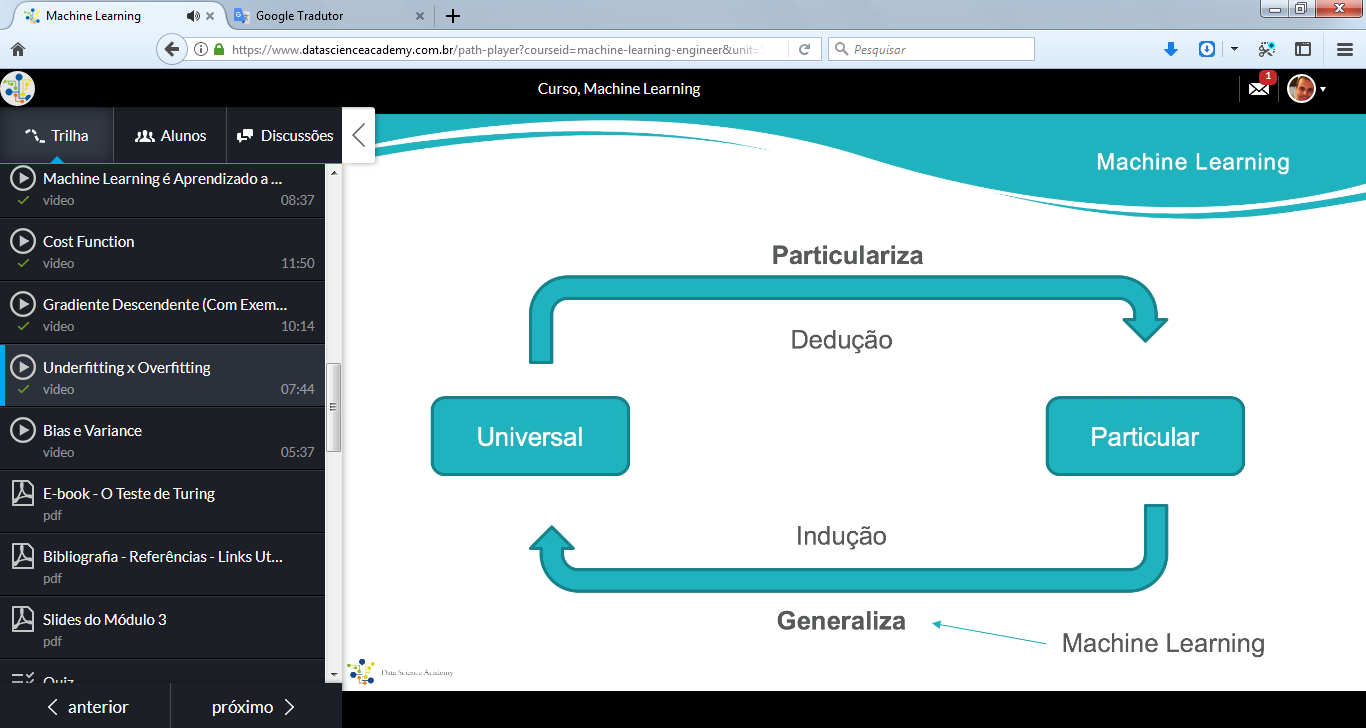
J -> Cost Funtion

**Gradiente Descendente**

Método que ajuda a minimizar o Cost Funtion, através de modificações de maneira iterativa.

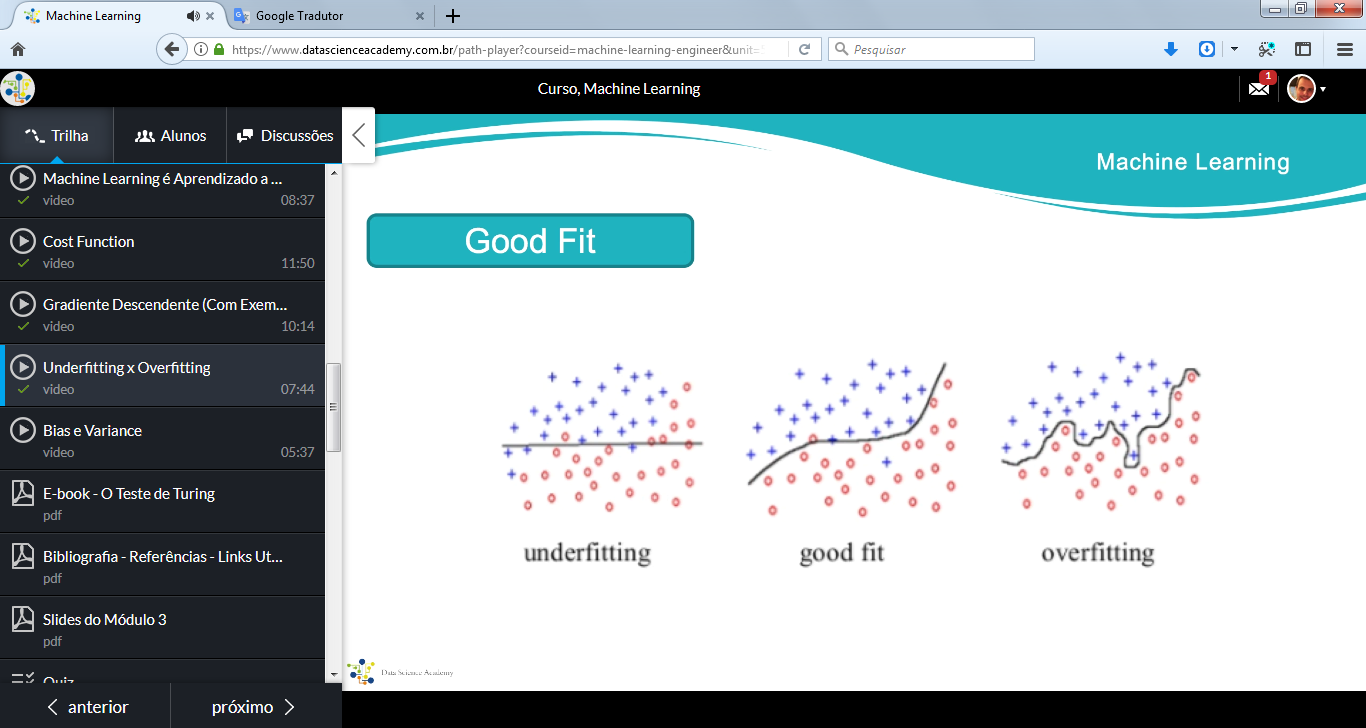


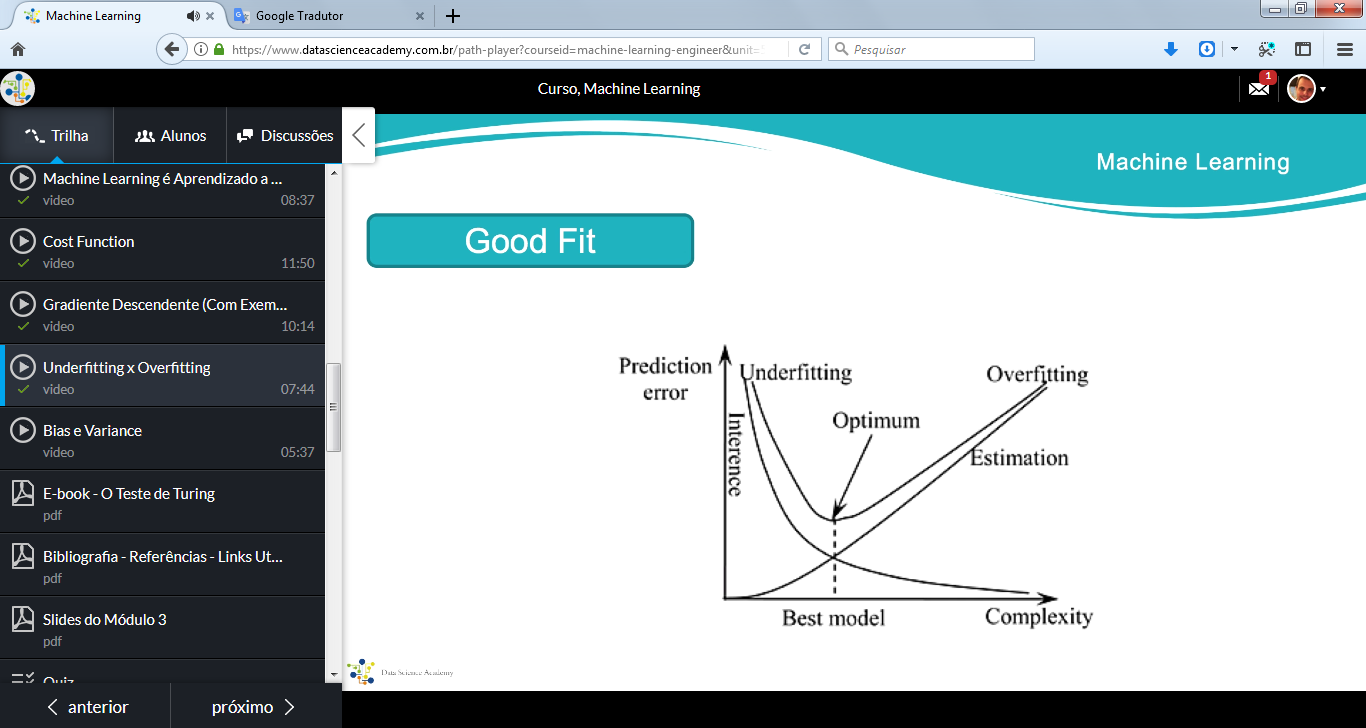
**Underfitting x Overfitting**



**Indução -** Definição de uma regra geral a partir de exemplos específicos

**Dedução -** Procura aprender conceitos específicos a partir de regras gerais.



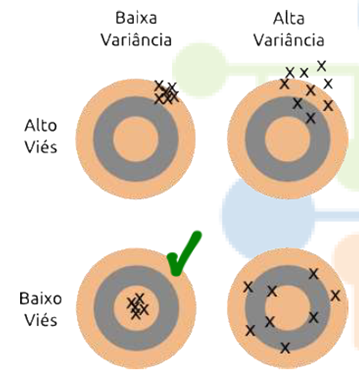


**Overfit -** Quando o modelo tem 100% no dataset de treino e 50% no dataset de test, enquanto deveria ter 75% em ambos os conjuntos.

**Viés (Bias) e Variância (Variance)**

Viés é a tendência do modelo aprender consistentemente uma generalização incorreta

Variância é a tendência de se aprender fatos aleatórios independentemente do sinal real



Ao escolhermos um modelo, nos deparando com o seguinte *tradeoff*:

**Baixo Viés, Alta Variância (Tipico caso de overfitting)**

Utilizar um modelo complexo que é capaz de reduzir consideravelmente o erro de previsão no dataset de treino, mas ao mesmo tempo não é tão generalizável a ponto de apresentar um bom resultado no dataset de teste.

**Alto Viés, Baixa Variância (Tipico caso de underfitting)**

Utilizar um modelo simples que é bem generalizável, mas não reduz consideravelmente o erro de previsão no dataset de treino.

A complexidade do modelo aumenta à medida que o número de variáveis preditoras aumenta.

