**Fundamentos de Machine Learning I**

Aprenda como os principais algoritmos de aprendizado de máquina funcionam, como treinar os modelos de classificação e avaliar seu desempenho.

**Nesta disciplina, você vai aprender os conceitos fundamentais dos algoritmos até a aplicação prática em problemas do mundo real.**

###### **Prof. Edneide**



**1. Planejamento**

## 1.1. O que você vai ensinar nessa aula?

## Qual habilidade, ferramenta, método ou conhecimento você vai ensinar nessa aula?

Nessa disciplina, o aluno vai aprender o conceito de aprendizado de máquina e como ele se diferencia do termo Inteligência Artificial. A partir dessa diferença, o aluno mergulha nos tipos de aprendizado e casos de uso da utilização de Machine Learning.

Como o conhecimento sobre os tipos de aprendizado, o aluno vai aprender o funcionamento e como realizar o treinamento dos principais algoritmos em cada tipo de aprendizado. Além de entender os indicadores de avaliação da performance e como aplicá-los usando a linguagem de programação Python, conceitos de garantia de aprendizado e a aplicação do funcionamento de cada algoritmo.

## 1.2. Na prática, o que o aluno será capaz de fazer após assistir a aula e aprender o conteúdo? (Ponto B)

Descreva o que o aluno será capaz de fazer após assistir a aula? Descreve as habilidades, ferramentas e métodos que ele vai aprender, como ele vai usar, em quais problemas e qual o resultado esperado que ele vai gerar no trabalho dele.

No final dessa disciplina, o aluno será capaz de treinar e medir o desempenho de vários algoritmos de Machine Learning de cada tipo de aprendizado. Além de técnicas para garantir que houve um real aprendizado do algoritmo e como o seu desempenho se comporta para cada variações dos seus parâmetros de controle.

## 1.3. Como o aluno resolve esse problema atualmente ? (Ponto A)

##### Descreve como o aluno resolve esse problema atualmente? Quais habilidades, ferramentas, métodos e conhecimento ele usa?

Atualmente, o aluno não conhece a família dos algoritmos de Machine Learning e portanto a sua contribuição na empresa fica limitada às análises descritiva e diagnóstica, já que não possui o domínio dos algoritmos capazes de entender comportamentos e destacar tendências passadas, para construir cenários futuros.

## 1.4. Quais os pré-requisitos dessa aula?

Defina as habilidades, ferramentas e conhecimentos necessários para o aluno entender perfeitamente o conteúdo da aula e conseguir aplicá-los no dia seguinte.

1. Requisitos computacionais:
   1. Windows 10
   2. Pelo menos 8GB de memória RAM
   3. Processador igual ou acima do i5
2. Requisitos conhecimento:
   1. Fundamentos da programação em Python



**2. Rascunho da Aula**

## 2.1. Brainstorm dos tópicos, habilidades, métodos ou conhecimento que serão ensinados nas aulas

##### Escreva todos os tópicos relacionados ao tema da disciplina que levarão o aluno do ponto A até o ponto B. Não se preocupe em colocar na ordem. Se preocupe somente em anotar TUDO o que você julga relevante ensinar. Nessa etapa, o professor realiza uma pesquisa bibliográfica para embasar sua disciplina.

### Ciclo 00: Introdução à Fundamentos de Machine Learning

##### Objetivo principal do ciclo: Ensinar os fundamentos da ferramenta

1. Visão geral da disciplina
2. Introdução aos Fundamentos de Machine Learning
3. O que é a Machine Learning
4. Instalação e configuração de um ambiente de desenvolvimento.
5. Exercícios de fixação

### Ciclo 01: Aprendizado de Máquina

##### Objetivo principal do ciclo: Organizar e limpar os dados

1. Machine Learning versus Inteligência Artificial
2. A utilização de Machine Learning
3. Casos de uso de Machine Learning
4. O aprendizado supervisionado
5. O aprendizado não supervisionado
6. O aprendizado por reforço
7. Exercícios

### Ciclo 02: As garantias do aprendizado

##### Objetivo principal do ciclo: Organizar e limpar os dados

1. Como garantir que o aprendizado aconteceu?
2. A estratégia de Treino-Teste
3. A estratégia de Validação Holdout ( Treino-Validação-Teste ) - Teoria
4. A estratégia de Validação Holdout ( Treino-Validação-Teste ) - Prática
5. O problema de Overfitting na Classificação
6. O problema de Underfitting na Classificação
7. As principais causas do Overfitting e Underfitting
8. Exercícios

### Ciclo 03: Algoritmos Bayesiano

##### Objetivo principal do ciclo: Organizar e limpar os dados

1. Como funciona o aprendizado supervisionado em classificação
2. Naive Bayes - Teoria
3. Naive Bayes - Prática
4. Naive Bayes - Exemplo de Uso
5. Métricas de avaliação I: Matriz de confusão
6. Métricas de avaliação II: Acurácia
7. Métricas de avaliação III: Recall
8. Métricas de avaliação III: Precision
9. Exercícios

### Ciclo 04: Algoritmos de Redes Neurais

##### Objetivo principal do ciclo: Organizar e limpar os dados

1. Introdução às redes neurais
2. A rede neural Perceptron - Teoria
3. A rede neural Perceptron - Prática
4. A rede neural Multi-Layer Perceptron - Teoria
5. A rede neural Multi-Layer Perceptron - Prática
6. Métricas de avaliação I - Log Loss
7. Exercícios

Dica para ensinar os algoritmos de Machine Learning:

Fase 1: Treinamento na prática

1. Treino - Validação - Teste
2. Treinamento do algoritmo.
3. Métrica de Treino, Métrica de Validação e a métrica de Teste

Fase 2: Processo de treinamento

1. Mostrar a receita de bolo do algoritmo.

Fase 2: Variação dos parâmetros

Fase 3: Dá indicações da ferramentas matemáticas que os alunos podem estudar.

### Ciclo 05: Algoritmos Random Forest

##### Objetivo principal do ciclo: Organizar e limpar os dados

1. Introdução a Random Forest
2. Algoritmo Random Forest - Treinamento ( Aplicação prática - Escrita de código )
3. Algoritmo Random Forest - Teoria ( Processo de treinamento )
4. Algoritmo Random Forest - Prática ( Variação dos valores )
5. Métricas de avaliação I - Curva Precision x Recall
6. Métricas de avaliação II - Curva ROC
7. Exercícios

### Ciclo 06: Algoritmo Extreme Gradient Boosting

##### Objetivo principal do ciclo: Organizar e limpar os dados

1. Introdução a Extreme Gradient Boosting (XGBoosting)
2. Algoritmo Extreme Gradient Boosting - Teoria
3. Algoritmo Extreme Gradient Boosting - Treinamento
4. Algoritmo Extreme Gradient Boosting - Prática
5. Métricas de avaliação I - F1-Score
6. Exercícios

### Ciclo 07: Algoritmos de Support Vector Machine

##### Objetivo principal do ciclo: Organizar e limpar os dados

1. Algoritmo Support Vector Machine - Introdução
2. Algoritmo Support Vector Machine - Teoria
3. Algoritmo Support Vector Machine - Treinamento
4. Algoritmo Support Vector Machine - Prática
5. Algoritmo Support Vector Machine - Regularização
6. Métricas de avaliação I - Kappa Score
7. Exercícios

### Ciclo 08: Projeto do Aluno

##### Objetivo principal do ciclo: Organizar e limpar os dados

1. Problema de negócio: Perguntas do CEO
2. O projeto do aluno
3. Prova de certificação.

Referência:

1. Machine Learning, An algorithm perspective, 2nd ( Stephen Marsland )
2. Python Machine Learning ( Machine Learning and Deep Learning with Python, sckit-learn, and tensorflow 2 ), 3rd ( Sebastian Raschka, Vahid MIrjalili )
3. An Introduction to Statistical Learning ( with Applications in R ), 2nd Edition



**3. Organização dos Ciclos**

Reúna as ferramentas, habilidades e conhecimentos necessários da sessão anterior para resolver um problema de negócio completo ou uma versão menor em cada ciclo. À medida que os ciclos avançam, a solução melhora com a adição de novas ferramentas e apresentações de conceitos mais aprofundados dos mesmos temas.

Ciclo 00: Introdução à Fundamentos de Machine Learning

Ciclo 01: Aprendizado de Máquina

Ciclo 02: As garantias do aprendizado

Ciclo 03: Algoritmos Bayesiano

Ciclo 04: Algoritmos de Redes Neurais

Ciclo 05: Algoritmos Random Forest

Ciclo 06: Algoritmo Extreme Gradient Boosting

Ciclo 07: Algoritmos de Support Vector Machine

Ciclo 08: Projeto do Aluno



# **4. Detalhamento dos ciclos**

## Para cada ciclo definido na sessão anterior, escreva os tópicos que serão abordados para resolver o problema de negócio proposto dentro de cada ciclo. Cada tópico é uma aula dentro do ciclo.

## Ciclo 00: Introdução a disciplina

O objetivo do ciclo 0 é conhecer a estrutura da disciplina, o professor, o método de avaliação e a projeto do aluno.

1. Introdução a micro formação em análise de dados
2. Os ciclos da disciplina: O que você vai aprender
3. A prova de certificação e o projeto do aluno
4. Como encontrar ajuda para dúvidas.

## Ciclo 01: XXX

O objetivo do ciclo 1 é conhecer o problema de negócio que o Power BI resolve, suas características e possibilidades. Além de instalar a ferramenta, ajustar as configurações iniciais e realizar um tour pelos principais componentes do painel.

**5. Projeto do Aluno**

Para praticar as habilidades e ferramentas adquiridas durante a disciplina, o aluno receberá um problema de negócio do mesmo nível, podendo ser ou não do mesmo tema e o produto final esperado que ele desenvolva.   
O aluno desenvolverá sozinho o PA, desde o entendimento do negócio até o produto final.

# **6.1 Problema de negócio do PA**

Faça a descrição do problema de negócio e coloque o link para conjunto de dados.

# **6.2 Produto de dados esperado**

Faça a descrição ou coloque o link do produto de dados final que o aluno vai construir ao finalizar o PA



**6. Prova de certificação**

# **7.1 Prova de múltipla escolha**

Escreva de 40 a 50 questões da prova de certificação. As perguntas podem variar entre perguntas teóricas com respostas diretas e perguntas com descrições pequenas de cenários ou problemas de negócio, perguntando o que o aluno faria ou qual decisão ele tomaria naquela cenário.

# **7. Descrição final da disciplina**

##### Nesta seção, você precisa descrever em linhas gerais:

##### Promessa ( O que a pessoa vai efetivamente aprender ? )

##### Problema ( Contextualizar o problema que originou a solução que o aluno vai aprender? )

##### Solução ( Breve descrição dos tópicos que o aluno vai aprender )

##### Conclusão ( O que o aluno será capaz de fazer quando ele terminar o curso todo )

Nesse curso, o aluno vai aprender as 4 habilidades fundamentais para qualquer profissional que deseja trabalhar com dados: SQL, Excel, Power BI e análise descritiva.

Profissionais que trabalham analisando dados ou criando relatórios têm dificuldades em fazer o seu trabalho de modo independente, necessitando do time de TI para fornecer os dados, ajuda com as fórmulas do Excel e uma explicação detalhada do passo a passo que é preciso fazer para analisar os dados.

Nessa micro formação, o aluno vai aprender a ferramenta de SQL para coletar os dados diretamente de um banco de dados e não depender mais do time de TI, vai aprender Excel para conseguir analisar uma planilha de dados sem ajuda e a ferramenta de Power BI para analisar os dados, criar indicadores e métricas para acompanhamento dos times. E por fim, o aluno vai aprender a escolher o tipo correto de análise que precisa realizar a partir de qualquer demanda do seu gestor, sem a necessidade de uma explicação detalhada do que precisa ser feito.

Ao terminar a micro formação em Análise de Dados, você vai parar de entregar relatórios e planilhas preenchidas e passará a entregar insights para tomada de decisão, levando você ao próximo passo profissional.

# **1.2. CONFIGURAÇÃO E INSTALAÇÕES DO COMPUTADOR**

##### Quais os requisitos mínimos de hardware e software para completar o curso.

* Computador com pelo menos 8GB de memória RAM.
* Processador i5 ou i7
* Sistema operacional Windows 10 Professional