DISCIPLINA....................... : Machine Learning para Negócios

SEMESTRE/ANO................. : 1º/2021

DEPARTAMENTO................ : TDS

**CURSO ........................... : CG**

**SEMESTRE CURRICULAR...... : DISCIPLINA ELETIVA, A PARTIR DO 4º SEMESTRE**

**NÚMERO DE CRÉDITOS....... : 2**

**PROFESSORES.................. : GUSTAVO MIRAPALHETA**

OBJETIVOs Da disciplina

Em praticamente todos os setores, as organizações enfrentam atualmente desafios analíticos sem precedentes. Os dados tornaram-se mais complexos, mais variados e cada vez menos estruturados. Neste contexto a análise e a tomada de decisão se apoiam cada vez mais em sistemas de inteligência artificial e aprendizado de máquina.

Esta disciplina concentra-se em como formular um problema de administração de empresas de forma que o mesmo possa ser resolvido por meio de técnicas de aprendizado de máquina, bem como no entendimento das soluções apresentadas pela máquina e na sua apresentação para o público decisor de alto nível.

A disciplina irá aborda técnicas de aprendizado de máquina supervisionado, tendo por base redes neurais. As aplicações principais serão: mineração de texto, processamento de linguagem, reconhecimento de imagens e de padrões em geral.

OBJETIVOs De aprendizagem

Os objetivos de aprendizagem da disciplina estão apresentados na tabela abaixo, demonstrando como os mesmos contribuem para os objetivos do CGA.

|  |  |  |
| --- | --- | --- |
| **Objetivos do CGA** | **Objetivos de aprendizagem da disciplina** | **Grau de contribuição** |
| Realidade do país |  | **○ ○ ○** |
| Contexto global |  | **○ ○ ○** |
| Conhecimento em administração |  | **○ ○ ○** |
| Solução de problemas | O aluno deverá ser capaz de identificar quais tipos de problema podem ou não ser resolvidos por meio de aprendizado de máquina e como os resultados propostos pela máquina podem ser implementados. | **● ● ○** |
| Ética |  | **○ ○ ○** |
| Diversidade |  | **○ ○ ○** |
| Comunicação escrita | O aluno deverá ser capaz de produzir relatórios e análises a partir das soluções geradas pela máquina. | **● ○ ○** |
| Comunicação multimídia |  | **○ ○ ○** |
| Métodos quantitativos | O aluno deverá ser capaz de modelar um contexo de negócios, identificando entidades e relacionamentos de forma que o problema possa ser resolvido por meio de aprendizado de máquina. | **● ● ●** |
| Raciocínio lógico | O aluno deverá ser capaz de elaborar pelo menos um modelo computacional aplicado a um problema específico de administração de empresas na área de machine learning. | **● ● ○** |

CONTEÚDO

1. Contexto atual de negócios: Big Data, Inteligência Artificial, Machine Learning
2. Ambientes de Desenvolvimento: Python Google/Colab
3. Fundamentos do aprendizado de máquina. Regressão linear, logística, support vector machiens e redes neurais.
4. Aplicações em administração de empresas: Mineração de Texto, Processamento de Linguagem. Reconhecimento de Imagens. Criatividade de máquina.

TÓPICOS PREVISTOS PARA AS AULAS

1. Cenário tecnológico atual. Python/Google Colab. Introdução ao aprendizado de máquina.
2. Métodos de previsão de valor. Regressão linear múltipla. Aplicações em finanças. Previsão do valor de ações.
3. Métodos de classificação: Regressão logística. Aplicações em modelos de crédito.
4. Seleção e validação de modelos. Aplicações em mineração de texto.
5. Técnicas avançadas de previsão e classificação: Random forest e Support vector machines.
6. Implementação de modelo preditivo em Python e scikit-learn.
7. Introdução às redes neurais e aos sistemas de deep learning. Bibliotecas tensorflow e keras. Fundamentos matemáticos do aprendizado de máquina com redes neurais: Backpropagation.
8. Deep Learning: Exemplos de aplicação de redes neurais em problemas de classificação e previsão.
9. Aplicações de negócios: redes neurais convolutivas: reconhecimento de imagens, recorrentes: processamento natural de linguagem, autogenerativas: criatividade de máquina.
10. Implementação de algoritmo de machine learning através da lógica de Qubits. (utilizando Qiskit, IBM)

Conforme o andamento do curso poderá haver adequações na sequência de aulas ou introdução de outros tópicos de interesse.

METODOLOGIA

Em aula os professores apresentarão e discutirão aspectos conceituais e os fundamentos das principais técnicas de análise, preparação de dados e de aprendizado de máquina. É fundamental que os alunos se aprofundem em cada tema através do estudo das referências e materiais indicados. A disciplina pressupõe que o aluno irá realizar atividades extra-classe regularmente (aproximadamente duas horas por semana, dependendo da facilidade do aluno para com o tema).

O aprendizado será alcançado com o uso intensivo da linguagem Python e do ambiente Google Colab. No entanto o conhecimento aprofundado destes ambientes não é pré-requisito para a disciplina, uma vez tanto a programação da mesma prevê uma etapa de nivelamento quanto espera-se do aluno a motivação para estudar e evoluir de forma rápida na familiaridade com estes ambientes.

CRITÉRIO DE AVALIAÇÃO

|  |  |
| --- | --- |
| **Notas** | **Peso** |
| 1ª Avaliação: Machine Learning (valor e categoria) | 30% |
| 2ª Avaliação: Deep Learning (categoria) | 35% |
| 3ª Avaliação: Projeto em Deep Learning. | 35% |

1. Em grupo. Art History of Florence
2. Em grupo. German Credit
3. Em grupo. Blu365

Bibliografia

Bibliografia básica:

CHOLLET, François *Deep Learning in Python.* Manning, 2018.

GARRETH, James *An Introduction to Statistical Learning with Applications in R.* Springer, 2013.

HILPISCH, Yves *Artificial Intelligence in Finance: A Python Based Guide.* O’Reilly, 2020.

Bibliografia complementar:

GERON, Aurelien. *Hands on Machine Learning*. O´Reilly, 2017

HASTIE, Trevor *The Elements of Statistical Learning.* Springer, 2009.

MEHTA, Nihal *Quantum Computing Program Next-Gen Computers for Hard, Real-World Applications.* Pragmatic Bookshelf, 2020.

SILGE, Julia *Text Mining in R.* O’Reilly, 2015

STRANG, Gilbert *Linear Algebra: Learning from Data.* Cambridge, 2019

WICKHAM, Hadley *Data Science in R.* O’Reilly, 2017

O aluno é incentivado a esclarecer suas dúvidas através de buscas na Internet e consultas a fórums de programação (exemplos: Stack Overflow e Quora).

Recursos online:

Caltech: <https://work.caltech.edu/telecourse> (Learning from Data)

MIT: <http://introtodeeplearning.com/> (Introduction to Deep Learning)

Stanford: <https://cs224d.stanford.edu/> (Natural Language Processing with Deep Learning)

Neural Net Playground: [https://playground.tensorflow.org](https://playground.tensorflow.org/) (Visualização de redes neurais)

COMPROMISSO ÉTICO – PROFESSOR/ALUNO

As aulas e atividades da disciplina serão pautadas pelo respeito mútuo, na sala de aula e fora dela. Haverá respeito do professor aos alunos, preparando-se para as aulas, respondendo a dúvidas, avaliando-os com isenção e imparcialidade e procurando ajudá-los no que estiver ao seu alcance. Haverá respeito por parte do aluno, ao professor e aos seus colegas, comportando-se de maneira ética e colaborando de forma significativa com o progresso das aulas, enriquecendo-as sempre que possível com perguntas e dúvidas pertinentes ao assunto abordado e preparando-se para as aulas de maneira a aproveitá-las ao máximo.

Apesar de haver farta disponibilidade de recursos para consulta na Internet, é considerado como compromisso ético nessa disciplina não copiar código de programas, mesmo de forma parcial, de outras pessoas ou fontes.

PROGRAMAÇÃO Aula-a-aula

A programação aula-a-aula é divulgada na primeira aula, pelo eClass.



CONTATO E *OFFICE HOURS*

|  |  |  |
| --- | --- | --- |
| Professor | Contato | Office Hours |
| Gustavo Mirapalheta | [gustavo.mirapalheta@fgv.br](mailto:gustavo.mirapalheta@fgv.br)  [gustavo.mirapalheta@gmail.com](mailto:gustavo.mirapalheta@gmail.com) | A combinar: whatsapp +55 11 9 9204 1201 |