



Bootcamp: Analista de Machine Learning

Plano de Ensino

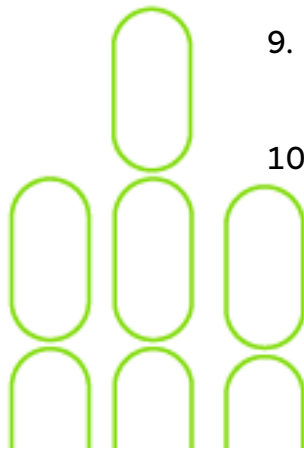
Módulo 2: Técnicas Avançadas Para Machine Learning

Descrição do Módulo

Este módulo tem como objetivo aprofundar o conhecimento dos alunos em técnicas avançadas de machine learning. Os alunos aprenderão sobre algoritmos avançados, como redes neurais artificiais, aprendizado por reforço e aprendizado de máquina com Spark. Além disso, eles serão introduzidos a técnicas de otimização de hiperparâmetros e validação cruzada.

Objetivo de Ensino

Espera-se que o aluno consiga, ao final deste Módulo:

1. Espera-se que o aluno consiga, ao final deste módulo:
 2. Compreender otimização de Hiperparâmetros de modelos;
 3. Compreender o uso de Validação Cruzada;
 4. Aprofundar em Ensemble Methods mais avançados;
 5. Aprender a aplicar soluções de Machine Learning em Spark;
 6. Compreender a metodologia de Aprendizado por Reforço;
 7. Aprender algoritmos de Redução de Dimensionalidade;
 8. Aprender algoritmos de Recomendação;
 9. Compreender os conceitos básicos por trás de Redes Neurais Artificiais.
 10. Implementar frameworks e melhores práticas de segurança de infraestruturas locais;
- 



Estrutura Analítica do Curso

Esta estrutura de tópicos deve conter os títulos das videoaulas, para que elas possam ser organizadas de maneira clara no Ambiente de Aprendizagem.

(Fique à vontade para alterar o número de capítulos e tópicos do modelo abaixo):

Módulo 2. Técnicas Avançadas Para Machine Learning

1ª Parte do Módulo

Capítulo 0 Visão geral do curso <videoaulas gravadas>

- 0.1. Apresentação do Professor
- 0.2. Apresentação do Módulo


Capítulo 1 Otimização de Hiperparâmetros <videoaulas gravadas>

- 1.1. Hiperparâmetros de um Modelo
- 1.2. Técnica Grid Search
- 1.3. Técnica Random Search

Capítulo 2 Validação Cruzada <videoaulas gravadas>

- 2.1. Introdução a Validação de Modelos
- 2.2. Métricas Mais Importantes
- 2.3. Leave-One-Out Cross-Validation (LOOCV)
- 2.4. K-fold Cross-Validation (KFCV)
- 2.5. Validação Cruzada em Dados Temporais

Capítulo 3 Ensemble Methods <e-book>

- 3.1. Vantagens e aplicações de Ensemble
 - 3.2. Bagging
 - 3.3. AdaBoost
 - 3.4. Gradient Boosting (GB)
 - 3.5. Stochastic Gradient Boosting (SGB)
 - 3.6. XGBoost
- 



Capítulo 4 Machine Learning com Spark <material complementar>

- 4.1. Introdução ao Spark
- 4.2. Biblioteca MLlib
- 4.3. Exemplo Prático: Caso de uso
- ...

Primeira Aula Interativa
Atividade de Fixação
2ª Parte do Módulo

Capítulo 5 Aprendizado por Reforço <e-book>

- 5.1. Introdução a Aprendizado por Reforço
- 5.2. Função de Utilidade
- 5.3. Valor Q^*
- 5.4. Exploração vs. Aproveitamento

Capítulo 6 Algoritmos de Redução de Dimensionalidade <videoaulas gravadas>

- 6.1. Introdução a Algoritmos de Redução de Dimensionalidade
- 6.2. PCA (Principal Component Analysis)
- 6.3. Transformação t-SNE

Capítulo 7 – Algoritmos de Recomendação <videoaulas gravadas>

- 7.1. N-NMF (Non-Negative Matrix Factorization)
- 7.2. Construindo um Sistema de Recomendação com N-NMF

Capítulo 8 Redes Neurais Artificiais <material complementar>

- 8.1. Introdução ao Aprendizado Profundo
- 8.2. Exemplo Prático: tensorflow playground

Segunda Aula Interativa



Atividade modular

Observação

Será considerado aprovado no Bootcamp e poderá obter o certificado de conclusão o aluno que:

- Atingir 70% de aproveitamento na soma total de pontos do Bootcamp.

Referências

AMERICAN National Standards Institute. The SQL Standard – ISO/IEC 9075:2016. Disponível em

<https://blog.ansi.org/2018/10/sql-standard-iso-iec-9075-2016-ansi-x3-135/> ;. Último acesso: 04, Abr 2020.

CHAMBERLIN, Donald D.; BOYCE, Raymond F. SEQUEL: A Structured English Query Language. Disponível em: <http://www.joakimdalby.dk/HTM/sequel.pdf> . Último acesso: 04, Abr 2020.

CHEN, Peter. Modelagem de Dados. São Paulo: McGraw-Hill, Makron, 1990.

CHEN, Peter. Peter Chen Home Page at Louisiana State University (LSU).

Disponível em: <https://www.csc.lsu.edu/~chen> . Último acesso: 01, Mai 2020.

CHEN, Peter. The Entity-Relationship Model - Toward a Unified View of Data. Disponível em:

<https://dspace.mit.edu/bitstream/handle/1721.1/47432/entityrelationshx00chen.pdf&> ,

Último acesso em 01 de Março de 2020.

CODD, Edgar F. A Relational Model of Data for Large Shared Data Banks.


Disponível em

<https://github.com/dmvaldman/library/blob/master/computer%20science/Codd%20-%20A%20Relational%20Model%20of%20Data%20for%20Large%20Shared%20Data%20Banks.pdf> . Último acesso em 03, Mar 2020.

CODD, Edgar F. **The Relational Model for Database Management: Version 2.** United States Of America: Addison Wesley Publishing Company, 1990.

CODD's Twelve Rules. Disponível em: <https://computing.derby.ac.uk/c/codds-twelve-rules> ;. Último acesso: 09, Mar. 2020.

COUGO, Paulo. **Modelagem Conceitual e Projeto de Banco de Dados.** 14ª reimpressão. Rio de Janeiro: Elsevier, 1997.



GARTNER Glossary. Disponível em <https://www.gartner.com/en/information-technology/glossary/big-data> ;. Último acesso: 02 de Maio 2021.

HILLS, Ted. **NOSQL and SQL Data Modeling: Bringing Together Data, Semantics, and Software**. Technics Publications, 2016.

IBM. A History and Evaluation of System R. Disponível em: <https://people.eecs.berkeley.edu/~brewer/cs262/SystemR.pdf> ;. Último acesso: 04, Abr 2020.

KORTH, Henry F. **Sistema de bancos de dados**. 3ª ed. São Paulo: McGraw-Hill, 1999.

MACHADO, Felipe Nery Rodrigues; ABREU, Maurício Pereira de. **Projeto de Banco de Dados: uma visão prática**. 13ª ed. São Paulo: Érica, 1996.

MCKINSEY Digital. Disponível em <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation> ;. Último acesso: 02 de Maio 2021.

MICHAELIS. Dicionário da Língua Portuguesa Brasileira. Disponível em: <https://michaelis.uol.com.br/moderno-portugues> ;. Último acesso: 05/04/2019.

MICROSOFT SQL Server 2019. Disponível em <https://www.microsoft.com/pt-br/sql-server/sql-server-2019> ;. Último acesso: 14, Out. 2022.

MICROSOFT SQL Documentation. Disponível em <https://docs.microsoft.com/en-us/sql/?view=sql-server-2019> ;. Último acesso: 14, Out. 2022.

MONQUEIRO, Julio Cesar Bessa. Programação Orientada a Objetos: uma introdução. Disponível em <https://www.hardware.com.br/artigos/programacao-orientada-objetos> . Último acesso: 05, Mar. 2019.

MONGODB. Site Oficial do Fabricante. Disponível em <https://www.mongodb.com> ;. Último acesso: 17, Out. 2020.

MONTEIRO, Leandro Pinho. Dados Estruturados e Não Estruturados. Universidade da Tecnologia, 2019. Disponível em <https://universidadetecnologia.com.br/dados-estruturados-e-nao-estruturados> ;. Último acesso: 02, Maio 2021.

NAVATHE, Shamkant B.; ELSMARI, Ramez. Sistemas de Banco de Dados. 4ª edição. São Paulo: Pearson Addison Wesley, 2005.

NOSQL Database Org. Disponível em: <http://nosql-database.org/> ;. Último acesso: 15, Mar. 2019.

PAT RESEARCH. Top 9 Object Databases. Disponível em <https://www.predictiveanalyticstoday.com/top-object-databases> . Último acesso: 05, Mar. 2020.

SIMPLILEARN. Introduction to NOSQL Databases Tutorial. Disponível em





<https://www.simplilearn.com/introduction-to-nosql-databases-tutorial-video> ;

Último acesso: 13, Mar. 2019.

STROZZI, Carlo. NoSQL: A non-SQL RDBMS. Disponível em

http://www.strozzi.it/cgi-bin/CSA/tw7/I/en_US/nosql/Home%20Page ;. Último

acesso: 11, Mar. 2019.

TAYLOR, Christiane. Structured vs. Unstructured Data. Disponível em:

<https://www.datamation.com/big-data/structured-vs-unstructured-data.html> ;

Último acesso: 04 de Maio 2021.

UNQL QUERY LANGUAGE UNVEILED BY COUCHBASE AND SQLITE. Disponível

em <https://www.couchbase.com/press-releases/unql-query-language> ;. Último

acesso: 11, Mar. 2019.

UNSQL Org. Disponível em <http://unql.sqlite.org/index.html/wiki?name=UnQL> ;.

Último acesso: 11, Mar. 2019.

