

# APRENDIZAGEM DE MÁQUINA

(usando Python)

Thiago Marzagão

VISÃO GERAL DA DISCIPLINA

- aplicações
- logística do curso
- preview do conteúdo

- recommender systems
- detecção de fraude
- inteligência artificial
- mercado financeiro
- lingüística

The Amazon logo, featuring the word "amazon" in a bold, black, lowercase sans-serif font. A yellow curved arrow starts under the letter 'a' and points to the right, ending under the letter 'n'.The Netflix logo, consisting of the word "NETFLIX" in a bold, white, uppercase sans-serif font with a black drop shadow. The text is centered on a solid red rectangular background.



# recommender systems

The screenshot shows the Amazon.com website interface. At the top, the browser address bar displays 'www.amazon.com' and the navigation menu includes 'Apps', 'news', 'programming', 'dataScience', 'posts', 'academia', 'random stuff', 'entertainment', 'CADE', and 'CÉREBRO'. The main heading is 'Recommendations for You, Thiago'. Below this, there are two rows of book covers. The first row includes: 'Statistical Analysis & Data Mining Applications', 'Learning pandas', 'Head First Java', 'Python Data Science Essentials', 'STATISTICS DONE WRONG', 'UNDERSTANDING MACHINE LEARNING', 'Test-Driven Development with Python', and 'Foundations of Machine Learning'. The second row includes: 'ALTIMETER THE BONDING AGENT', 'SQL 10', 'SQL Cookbook', 'Learning SQL', 'SQL 10', 'Python', 'DATA SCIENCE FROM SCRATCH', 'DATA SCIENCE FROM SCRATCH', 'Python Machine Learning', 'DATA SCIENCE FROM SCRATCH', and 'DATA MINING'. Below the second row is a link 'See all your recently viewed items >'. At the bottom of the recommendations section is a 'Back to top' link.

Get to Know Us  
Waiting for www.amazon.com...

Make Money with Us

Amazon Payment Products

Let Us Help You

# recommender systems

The screenshot shows the Netflix website interface for a user named Thiago. The browser address bar displays `www.netflix.com/browse`. The navigation bar includes the Netflix logo, a "Browse" dropdown menu, and "DVD" text. On the right, there is a search icon, a notification bell, and the user's name "Thiago" with a dropdown arrow. The main content area is divided into three sections:

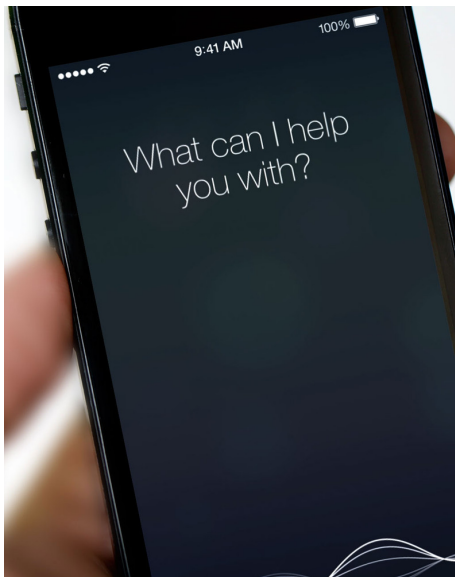
- Top Picks for Thiago:** A row of six movie and TV show thumbnails. From left to right: "12 MONKEYS" (a red background with a black monkey face), "NETFLIX JESSICA JONES" (a purple background with a woman's face), "TWO MEN UNDER ONE ROOF" (a man and a woman in a room), "THE WORLD OF BRAD PITT" (a street scene with a blue tint), "BROOKLYN NINE-NINE" (a yellow background with a police officer), and "MOCKINGJAY PART 1" (a fiery background with a woman's face).
- Trending Now:** A row of seven thumbnails. From left to right: "ORANGE IS THE NEW BLACK" (a black background with white text), "NETFLIX NARCOS" (a man's face), "how I met your mother" (a group of people), "SUITS" (a man in a suit), "RuPaul: DRAG RACE" (RuPaul's face), and "SONS OF ANARCHY" (a man's face).
- Watch It Again:** A row of seven thumbnails. From left to right: "Breaking Bad" (a man's face with green smoke), "Family Guy" (the cartoon characters), "NETFLIX HOUSE OF CARDS" (two men in suits), "WALKING DEAD" (a city street), "FRIENDS" (the main cast), and "SHERLOCK" (Sherlock Holmes).

At the bottom left, there is a small status message: "Waiting for cdn0.nflximg.net...". At the bottom right, there are navigation icons for back, forward, and search.

## detecção de fraude



# inteligência artificial



# inteligência artificial



# mercado financeiro

bdm.unb.br/handle/10483/10732

Apps news programming dataScience posts academia random stuff entertainment CADE CÉREBRO

**BDM**  
Biblioteca Digital de Monografias

Página inicial | Comunidades e Coleções | Data | Autor | Orientador | Título | Assunto | Tipo de documento | Ajuda

BDM >  
[Monografias de Graduação](#) >  
[Administração](#) >

Please use this identifier to cite or link to this item: <http://bdm.unb.br/handle/10483/10732>

File	Description	Size	Format
<a href="#">2014_SarahSabinodeFreitasMarcelino.pdf</a>		1,8 MB	Adobe PDF <a href="#">View/Open</a>

**Título:** Formação de portfólio por meio de Máquinas De Suporte Vetorial

**Autor(es):** [Marcelino, Sarah Sabino de Freitas](#)

**Orientador(es):** [Albuquerque, Pedro Henrique Melo](#)

**Assunto:** [Máquinas de Suporte Vetorial](#)  
[Portfólios \(Ações\) - formação](#)  
[Benchmarking \(Administração\)](#)

**Data de apresentação:** 20-Nov-2014

**Data de publicação:** 28-Jul-2015

**Referência:** MARCELINO, Sarah Sabino de Freitas. Formação de portfólio por meio de Máquinas De Suporte Vetorial. 2014. 98 f., il. Monografia (Bacharelado em Administração)— Universidade de Brasília, Brasília, 2014.

**Resumo:** A presente pesquisa teve como objetivo replicar a metodologia de Máquinas de Suporte Vetorial proposta por Fan e Palaniswami (2001) no contexto brasileiro de formação de portfólio. O SVM foi então utilizado para verificar se o uso de Máquinas de Suporte Vetorial na formação de portfólios de fato contribui para que o retorno seja superior ao de um benchmark do mercado, sendo que o ativo escolhido para tal comparação foi o



## preview do conteúdo

- Python
- regressão linear
- classificação
  - árvores de decisão
  - máquinas de suporte vetorial (SVM)
- clusterização
- pré-processamento
- mineração de textos
- raspagem de sites (webscraping)



# aprendizagem de máquina vs econometria

- Econometria: foco em *explicar* o mundo.
- Aprendizagem de máquina: foco em *prever* o mundo.
- Perguntas centrais em econometria:  $x$  e  $y$  realmente estão relacionados? qual a magnitude dessa relação?
- Pergunta central em mineração de dados: dados  $x$ , qual o valor esperado de  $y$ ?

## aprendizagem de máquina vs econometria

- Econometria: usa 100% dos dados p/ treinamento.
- Aprendizagem de máquina: usa 70%-99% dos dados p/ treinamento.
- Lógica:
  - ... estimar modelo usando dados de treinamento
  - ... testar modelo usando dados de teste
- Fundamental: dados de teste NÃO devem ser usados p/ *estimar* o modelo, apenas p/ *testar* o modelo.
- Objetivo é testar o modelo com dados ainda não vistos.

# LOGÍSTICA DO CURSO

- site do curso: <http://thiagomarzagao.com/teaching/ipea>
- no site: programa, slides, datasets, links p/ leituras
- comunicação: Slack (<https://ipea-ml.slack.com>)
- avaliação: trabalho final, individual