

MINERAÇÃO DE DADOS

Thiago Marzagão¹

¹ marzagao.1@osu.edu

VISÃO GERAL DA DISCIPLINA

ROTEIRO

- aplicações
- mercado de trabalho
- logística do curso
- preview do conteúdo

APLICAÇÕES

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

RECOMMENDER SYSTEMS



RECOMMENDER SYSTEMS

www.amazon.com

Apps news programming dataScience posts academia random stuff entertainment CADE CÉREBRO

Recommendations for You, Thiago

See all your recently viewed items >

Back to top

Get to Know Us

Waiting for www.amazon.com...

Make Money with Us

Amazon Payment Products

Let Us Help You

RECOMMENDER SYSTEMS

www.netflix.com/browse

Apps news programming dataScience posts academia random stuff entertainment CADE CÉREBRO

NETFLIX Browse DVD Search Thiago

Top Picks for Thiago

Trending Now

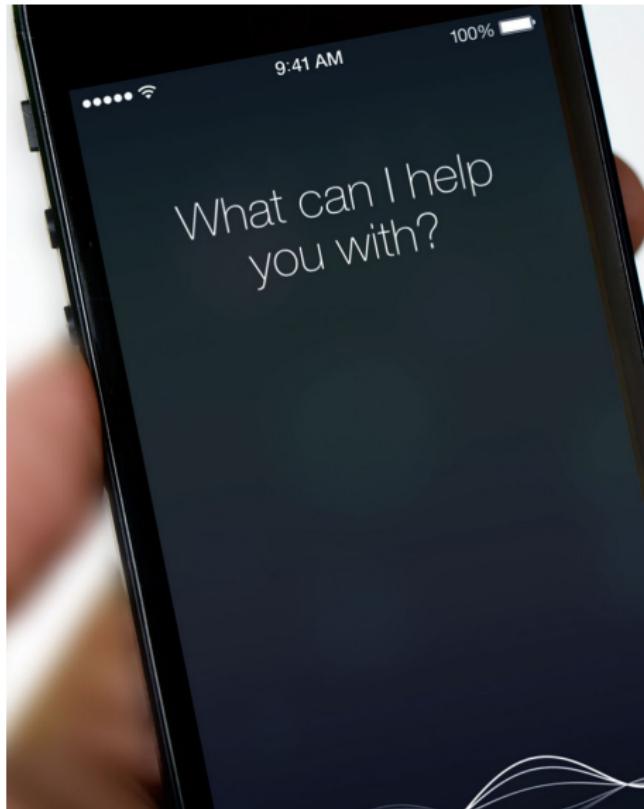
Watch It Again

Waiting for cdn0.netfliximg.net...

DETECÇÃO DE FRAUDE



INTELIGÊNCIA ARTIFICIAL



INTELIGÊNCIA ARTIFICIAL



MERCADO FINANCEIRO

← → C 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>

Files in This Item:			
File	Description	Size	Format
2014_SarahSabinodeFreitasMarcelino.pdf		1,8 MB	Adobe PDF View/Open

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

LINGÜÍSTICA

thing harem hear take never them kind are who dah wolves need line got yes
number went mary note going emsp speak gone madam drinkin' judge
patient hain't dollars three mumps word 'issel
everybody town harker's godalming live like loose along he's king
wednesday anyhow anyway afraid tom like
lady loving pretty out shall john wolf
maybe george child tom harker
suppose anything de than harker blow
yours why ear de than john boy
reckone raft mina around augst
what's breakfast helsing you've professor
an' art came september am's
once diary else e' buck jonathan till count
kne found seward didn't assured journal confidence
stolen evening right injured farewell certainly days yes'm biggest
right seward has master horror great something
mind breast small does mum sweet ole will
has must bring not san hello won't
sbreast boxes or cairn wreck bottom read squirm
mind huck alas nation wreake
care make fifteen
said stay may

MERCADO DE TRABALHO

www.kdnuggets.com/2013/02/salary-analytics-data-mining-data-science-professionals.html

Apps news programming dataScience posts academia random stuff entertainment CADE CÉREBRO

KDnuggets™ Data Mining, Analytics, Big Data, and Data Science Subscribe to **KDnuggets News** | Follow @KDnuggets | Contact search KDnuggets Search

Data Mining Software | News | Jobs | Academic | Companies | Courses | Datasets | Data Mining Course | Education | Meetings | Polls | Webcasts

Most Viewed

- TensorFlow Disappoints - Google Deep Learning falls shallow.
- 7 Steps to Mastering Machine Learning With Python.
- 5 Best Machine Learning APIs for Data Science
- R vs Python for Data Science: The Winner is ...
- A Statistical View of Deep Learning
- Amazon Top 20 Books in Databases & Big Data.
- 9 Must-Have Skills You Need to Become a Data Scientist

Most Shared

- The different data science roles in the industry

Table 1: 2013 Annual Income/Salary by Employment type

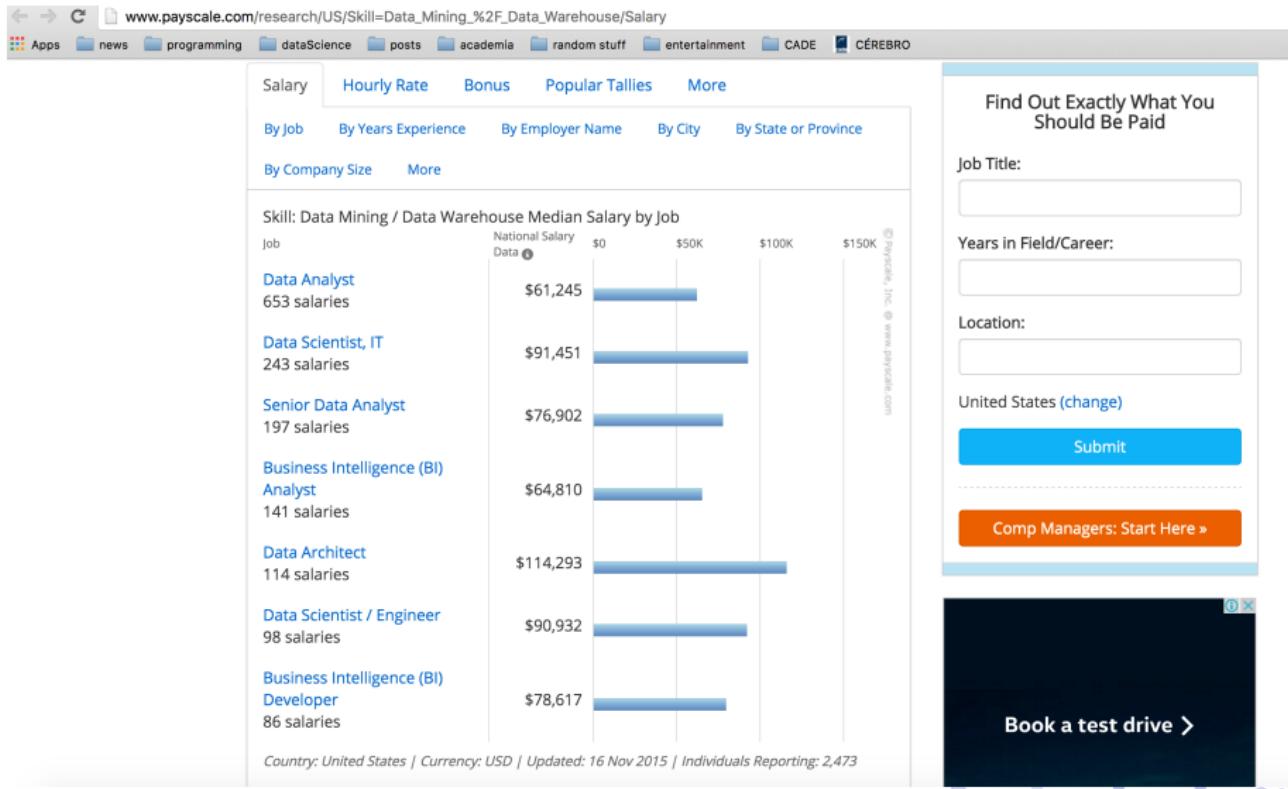
Employment	2013 Avg. Salary	2012 Avg. Salary	% Change	2013 Count	2012 Count
Self-employed	136.4	105.8	29.0%	25	26
Company	111.3	101.6	9.5%	284	375
University/Academia	89.6	63.2	41.9%	42	49
Government	72.5	71.8	1.0%	12	14
Student	33.4	34.1	-2.0%	19	17
Unemployed/Retired	20.0	48.3	na	1	6
All	105.2	94.1	11.9%	383	487

For estimating average salary we used mid-point of the range, 20K for * \$180" range.

We note also that there was only 1 self-reported Unemployed/Retired, vs 6 in 2012.

In Table 2 we break down the data by region. Since the numbers of self-employed and government data scientists are relatively small, we grouped them with industry and university, respectively. Students and unemployed were excluded.

MERCADO DE TRABALHO



PREVIEW DO CONTEÚDO

- algoritmos de mineração
 - regressão linear
 - classificação
 - árvores de decisão
 - máquinas de suporte vetorial (SVM)
 - clusterização
 - mineração de textos

PREVIEW DO CONTEÚDO

- computação
 - programação
 - Python
 - SQL
 - computação em nuvem
 - trabalhando c/ dados sujos

LOGÍSTICA DO CURSO

- site do curso: <http://thiagomarzagao.com/teaching>
- no site: programa, slides, datasets, links p/ leituras obrigatórias
- diferente de outra matérias: capítulos indicados devem ser lidos *após* a aula

LOGÍSTICA DO CURSO

- avaliação:
 - 5 competições, de igual peso
 - quanto melhor a previsão do modelo, melhor a nota