

Data Science e Machine Learning com Python

Gileno Alves Santa Cruz Filho
@GilenoFilho

Gileno, quem?

- Interesse em: Desenvolvimento Web e Científico, Inteligência Artificial, Engenharia de Avaliações, Design, Usabilidade, Minimalismo e Empreendedorismo





PyCursos



PYTHON PARA ZUMBIS



CURSO
GRÁTIS



**DANTAS ENGENHARIA
DE AVALIAÇÕES**

BIG DATA

RESULTS
SOCIAL YEARS
DRIVE SAN NAS
REPLICATION SOURCES
HANDLE SOFTWARE
COMPLEX CHALLENGES
PAST ANNOUNCED LIMITS
LARGE STORAGE
HAND CAPACITY
MILLION SETS
INCLUDE NETWORKS
HYPOTHESIS SCIENCE
END NEARLY MAPREDUCE
FUTURE MOST PARALLEL STATE CRITIQUE
CRITIQUES APPROACH MANY
OUT DEVELOPMENT TECHNOLOGY
BILLION APPROACH NEW
CURRENTLY PEOPLE TRAFFIC
TERABYTES RELATED TIME
INITIATIVE MARKET VARIETY
ECONOMIC SENSOR
PROJECT ANNUAL NOW
DISTRIBUTED RELEVANT
PER USED

GOVERNMENT
FACTOR SYSTEMS MANAGE ONLINE DAY
MAKING INTEGRATION WORLDWIDE
VOLUME
PROCESS SIZE
MASSIVE DEFINITION BETWEEN STATISTICS
BUSINESS EFFECTIVE VISUALIZATION
DATABASES PROCESSING ANALYTICS
SIMULATIONS

TECHNOLOGIES
USING EXabytes DETERMINE WELL
SET DEPARTMENT QUERIES HIGH USERS
HIGHER ALGORITHMS COMPANIES
APPLICATIONS REAL RATE STRUCTURE WORK
INFRASTRUCTURE

MANAGEMENT
PRIVACY INTERNATIONAL IMPLEMENTATION BECOMES
SECOND NATIONAL INTELLIGENCE
FLOW APPROACHES
INTERNET ARCHITECTURE

PETABYTES
ANALYSIS TERMS
PRIVATE RESEARCH
LEARNING LESS HUGE
EVERY WORLD VARIOUS
BASED TIMES
PROCESSES APPROACHES
PARADIGM FUNDING
RECORDS COST INSIGHT TOOLS
CENTER

Data Science

BUZZWORDS

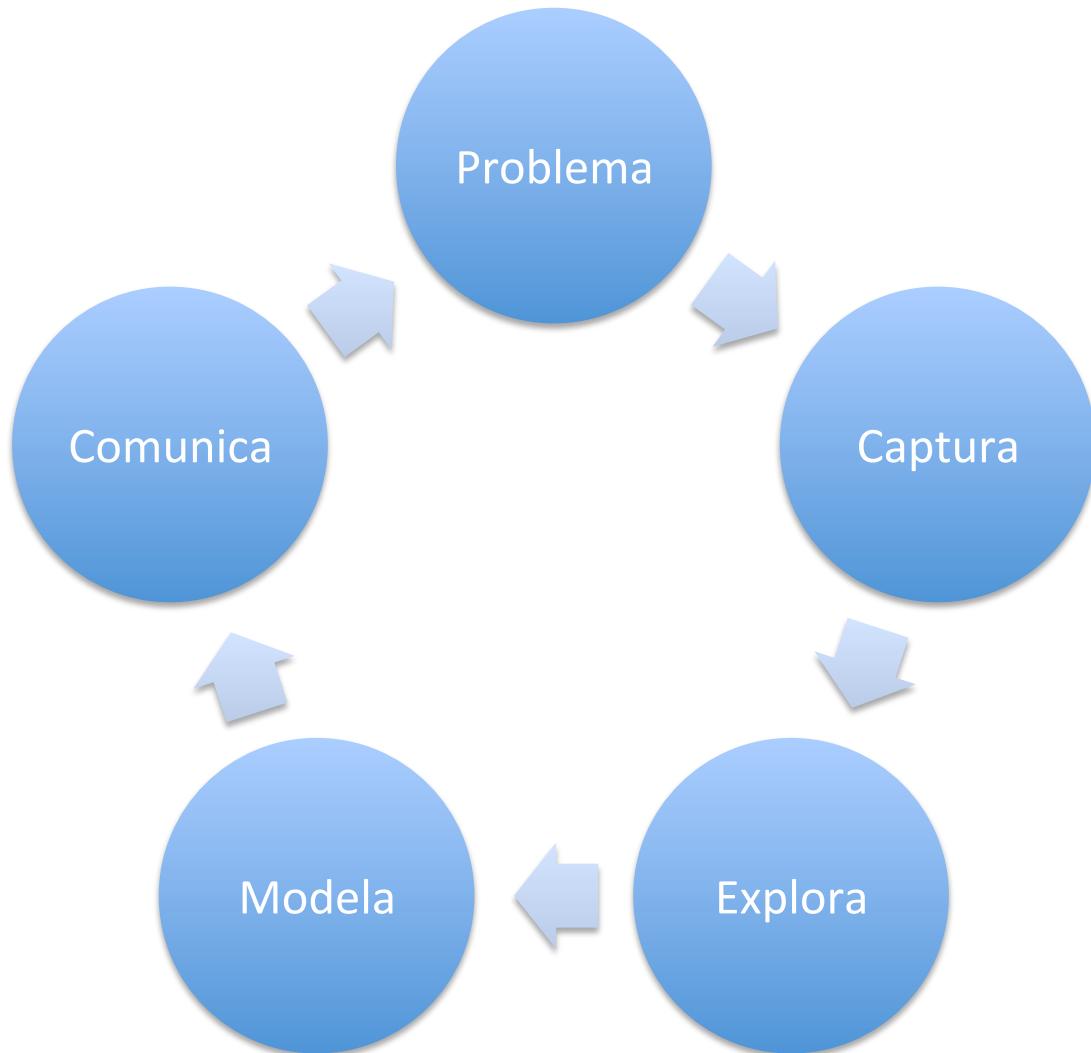


BUZZWORDS EVERYWHERE

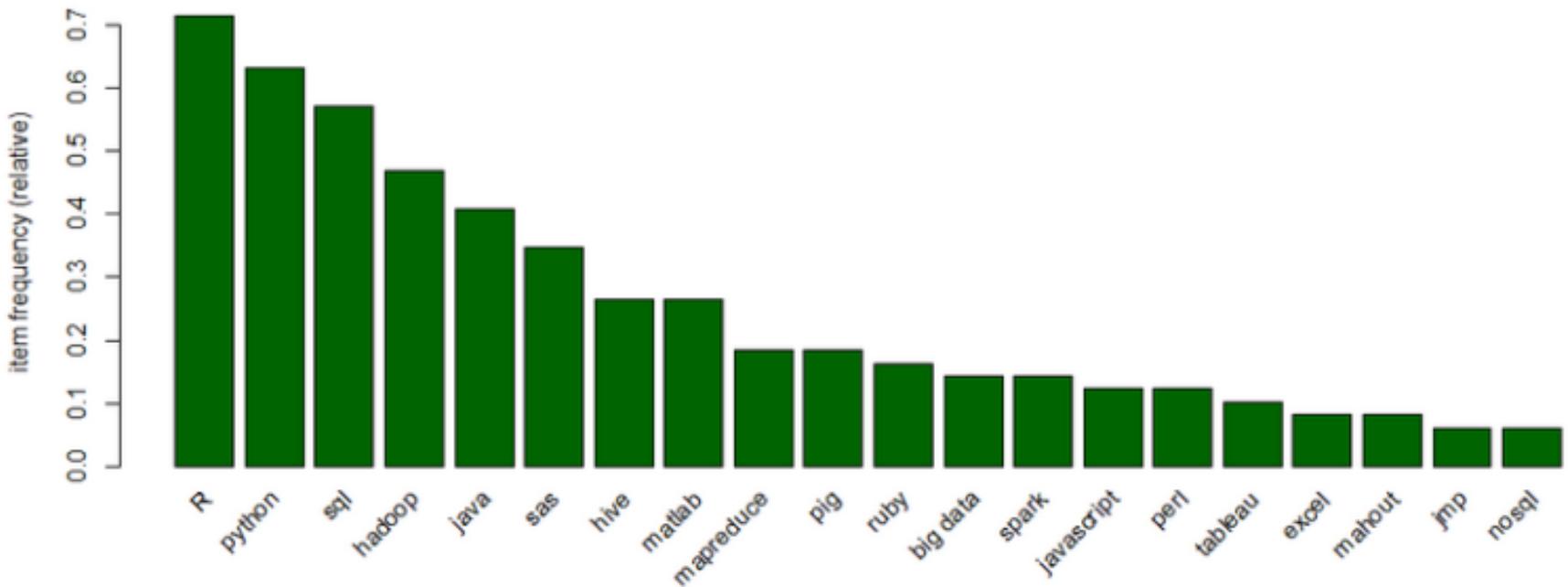
Data Science



Data Science



Porque Python?



<http://www.datasciencecentral.com/profiles/blogs/popular-software-skills-in-data-science-job-postings>

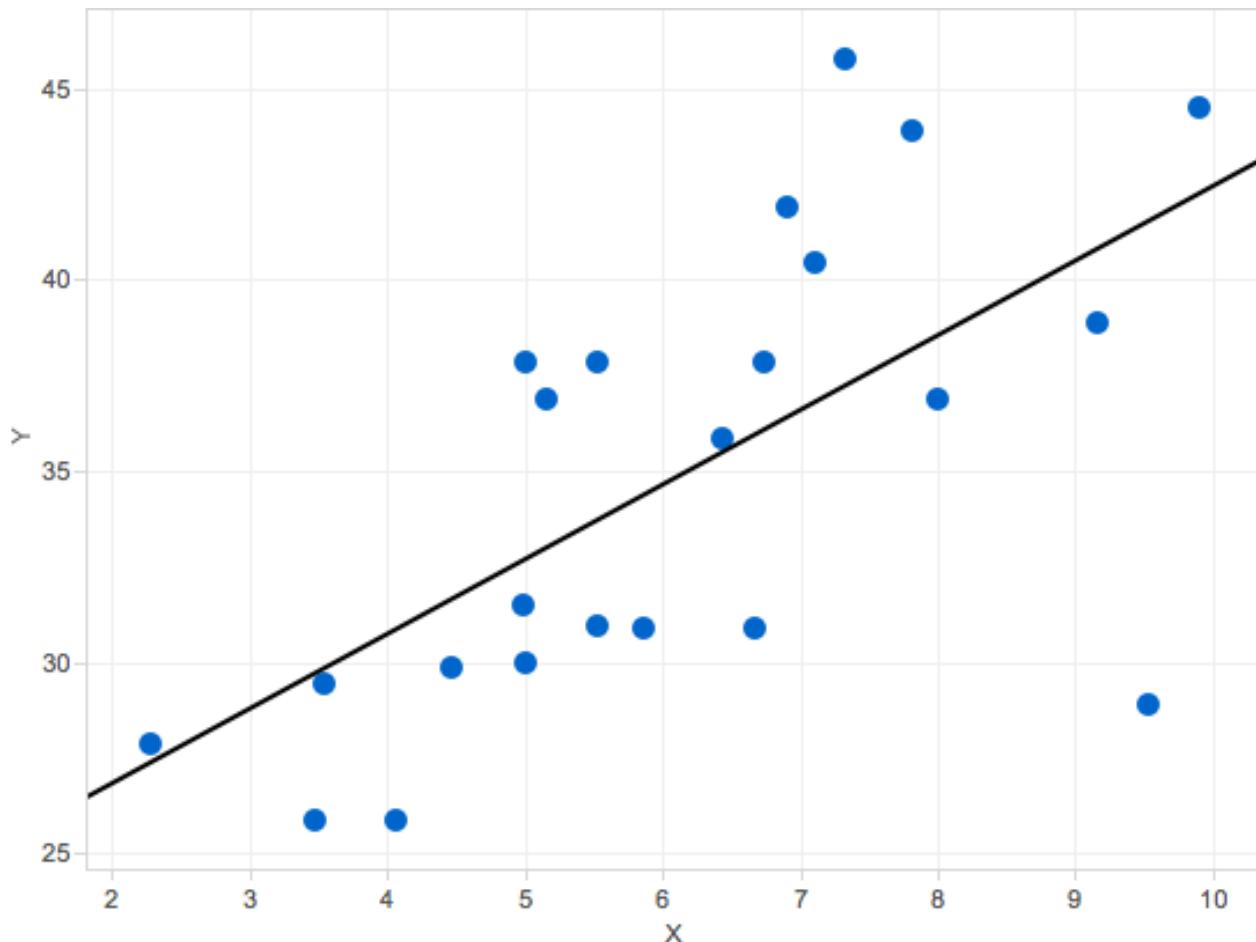
Que problemas posso resolver?

- Classificar se um e-mail é spam ou não
- Descobrir o valor de um imóvel
- Recomendar um produto a um usuário
- Descobrir padrões e agrupamentos
- Reconhecimento de faces/identidades em imagens ou vídeos

CODING

- Vamos brincar um pouco:
- Todo código em: <https://github.com/gileno/poticon>

Regressão



Régressão

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Árvore de Decisão

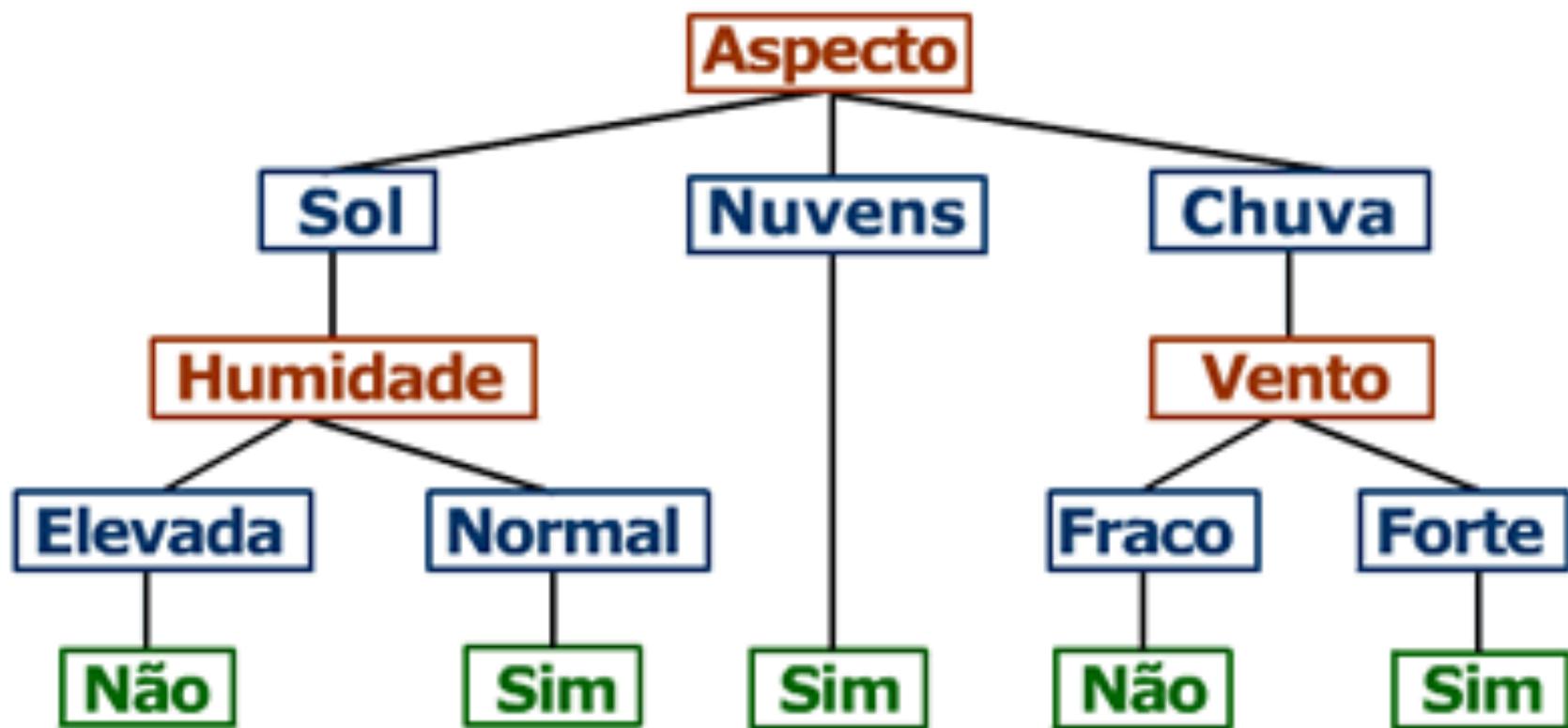
Exemplos de Treino

Dia	Aspecto	Temp.	Humidade	Vento	Jogar Ténis
D1	Sol	Quente	Elevada	Fraco	Não
D2	Sol	Quente	Elevada	Forte	Não
D3	Nuvens	Quente	Elevada	Fraco	Sim
D4	Chuva	Amenos	Elevada	Fraco	Sim
D5	Chuva	Fresco	Normal	Fraco	Sim
D6	Chuva	Fresco	Normal	Forte	Não
D7	Nuvens	Fresco	Normal	Fraco	Sim
D8	Sol	Amenos	Elevada	Fraco	Não
D9	Sol	Fresco	Normal	Fraco	Sim
D10	Chuva	Amenos	Normal	Forte	Sim
D11	Sol	Amenos	Normal	Forte	Sim
D12	Nuvens	Amenos	Elevada	Forte	Sim
D13	Nuvens	Quente	Normal	Fraco	Sim
D14	Chuva	Amenos	Elevada	Forte	Não



Árvore de Decisão

Árvore de Decisão para Jogar Ténis



<http://pycursos.com/pandas/>





Links

- Google