

Inteligencia artificial avanzada para la ciencia de datos I

TC3006C.71 Módulo 5 NLP 1

EMMANUEL PÁEZ

emmanuel.paez@tec.mx



EXPERIENCE

ITC 2017
MSC Artificial Intelligence UK 2019
Consultant and instructor Google
Cloud Platform 6 years

CONTACT

emmanuel.paez@tec.mx
CSF and CCM



GCP CERTIFICATIONS



ABOUT ME

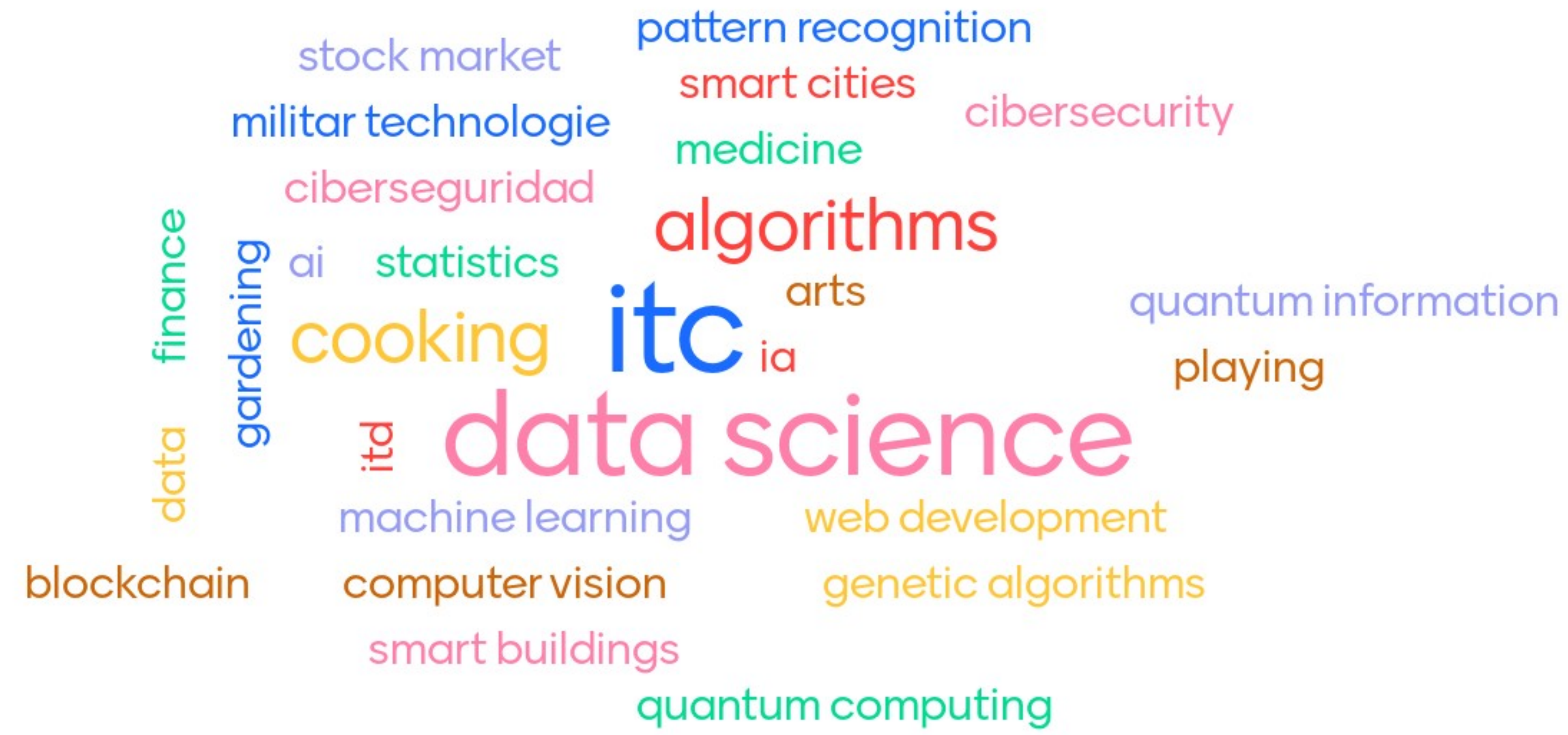
I love Artificial Intelligence,
Natural Language Processing,
Disruptive Technologies and
Data Science.

Science Fiction Fan.

Instructions



Type your career / interest areas
39 Responses



What do you expect from this course?

43 Responses



What do you think NLP means?

14 Responses

Has to do something with a search engine? Or probabilistic pattern unification of words.

Recognize patterns on language

Understanding and outputting human language

Language Understanding and generation

Making easier for a computer to understand humans and viceversa

Replicate human thinking process

Emulation of a real human response.

Natural language processin

semi supervised learning

What do you think NLP means?

14 Responses

Has to do with to simulate the way humans express or communicate

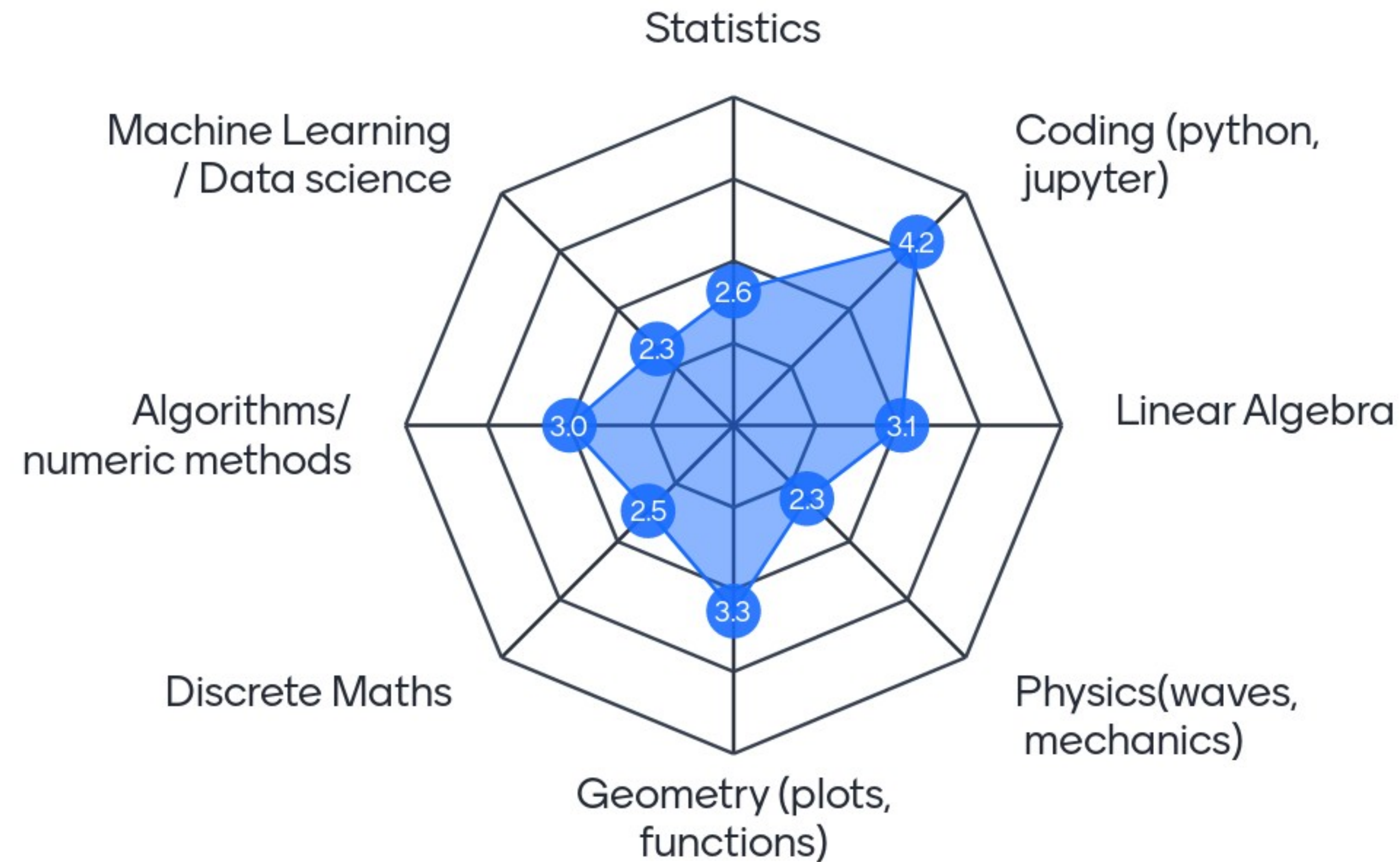
Emulation of human communication interactions

Natural language processing

NLP stands for Natural Language Processing. Considering the previous definition, we can spect that NLP enables computers to understand in a better way how humans communicate for better computing resu

Trying to extract sentiments, ideas, etc. from human language

Choose your confidence level on each area



Evaluation

- Weekly Quizzes (each friday, individually) 5% total
- Notebook evaluation (at the end, by teams) 5% total
- Evaluation(at the end, individually) 2% total
- Weekly quizzes + Notebook +Evaluation = 12% (Module 5)

Actividades	Ponderacion
Module 1. Python Libraries	10%
Module 2. Statistics for Data Science	10%
Module 3. Machine Learning	10%
Module 4 Big Data and AWS	10%
Module 5. Natural Language Processing 1	10%
Evaluation Module	10%
Total	60%

Evidencia	Ponderacion
Presentación argumentativa	40%

Index 1st part

1. Sound and Waveforms
2. Sound characteristics
3. Analog and digital Signals
4. Time Domain Features for ML

Index 2nd part

1. Corpus creation
2. Tokenizers /N-grams
3. PCA
4. Word2vec
5. Deep Learning
6. Transformers

Notes

Wednesday, August 30th class will be virtual

Friday, August 1st class to be defined

Quiz will be available

