

amplemarket

Machine learning!

(but in the real world)

29-06-22 / LIP Symposium - Coimbra
Duarte OC. & Olavo B.

1. Who are we?



Machine Learning Engineer

Duarte O. Carmo

Originally from Portugal,
now living in Denmark

Background in Strategy,
Consulting, Ventures

Enjoys running, coding,
reading (and beer)



Machine Learning Engineer

Olavo Bacelar

From Porto

Background in Statistics,
Physics, and... Statistical
Physics

Loves hiking and learning
all kinds of things

2. What 's Amplemarket?



amplemarket

A new playbook for sales

Transform how B2B companies do sales and grow

Leverage AI to build sales tools that support teams doing sales

Help *our* customers find *their* right customers

We've been growing since 2014 and just received 12 mUSD



Y-Combinator
investment

Amplemarket
as a Service

Amplemarket
as a Product

++ Investors
& growth

\$12 mUSD
series A

...maybe you
join?

2014

2016

2019

2020

2021

2022

Previously
Orankl

ORANKL

Our founding team is 100% NERD

Former LIP researchers

Co-founder

Luis Batalha

Luis' focus is on Product, ML and People. He helps drive Product development and works closely with the People team to expand Amplemarket worldwide.

NERD



Co-founder and CEO

João Batalha

Focus on Engineering. If there's anything that our Engineering team can't answer, he's your guy.

NERD

Co-founder

Mica Oliveira

Mica's main focus is Customers. He works closely with the Marketing, Growth, Sales and Success teams to drive strategy for activation, revenue and retention.

NERD

We're 100% remote and as distributed as they come

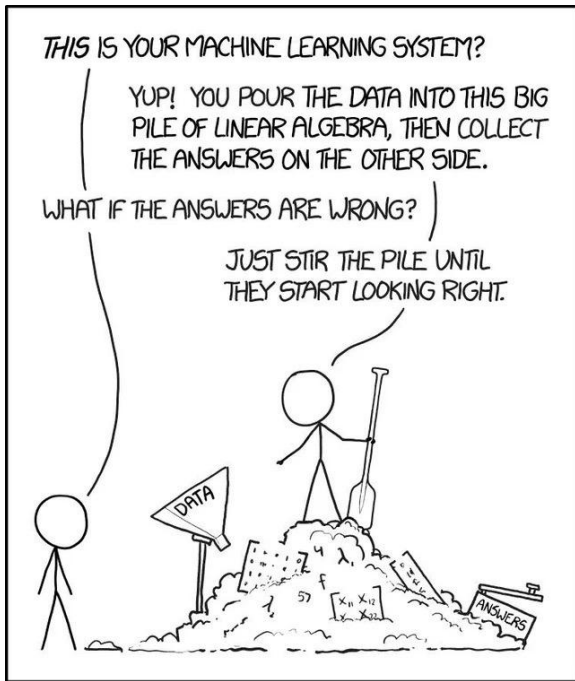


3. How do we do things?

**“WE NEED A MACHINE
LEARNING MODEL”**

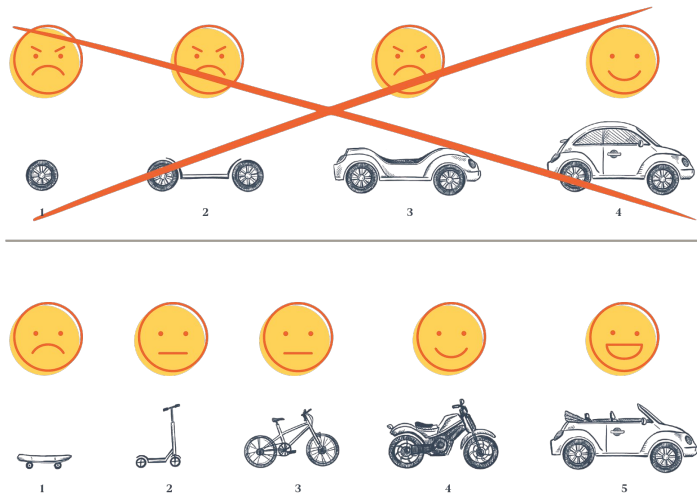
(you probably don't)

We don't start with Machine learning,
we start with people



- Define the **business goal**, and the **success metric**
- This is real world **(bad) data** - not Kaggle: cr*p in, cr*p out
- **Start with heuristics**, and increase complexity as needed
- **Put it out there** as fast as possible, then iterate

Our focus is to do applied research that directly improves our users' experiences



- We don't spend too much **time in the basement**
- Incredible models are **useless if not shared** with users
- Best model \neq best solution for the users/business (business metric)
- Quick iterations guarantee you are solving the **right problem**

We're makers at heart, and treat our schedules like it

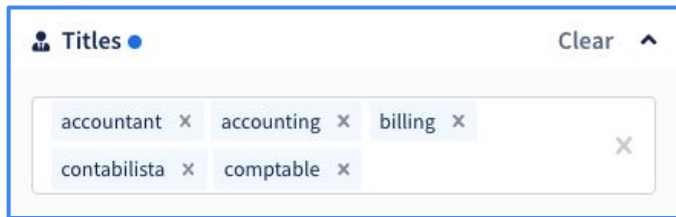


- Minimize time in meetings and double down on communication
- Fridays = no meetings
- We are on an emerging tech field, studying is important
- We are builders of things, disruptions are not welcome

4.1.Job title classification

Job titles help you find good leads, but there are millions of them

- 38 million different titles in our database
- Many ways to search for relevant people:



- **GOAL:** Categorize job titles into multiple functions or departments
- Titles are not easy:

"Java Ninja"

*"Developer
advocate"*

"Digital overlord"

"Co-Founder and Board Member"

We started with a typical machine learning approach

- Find dataset that maps titles to categories
- Quality check and clean data
- Train model, e.g: *fastText*
- Some issues:
 - Small dataset
 - Single-label
 - Lots of bad labels

Accounting	78.3%	0.0%	0.3%	1.5%	0.3%	10.0%	0.5%	0.0%	0.0%	3.1%	0.5%
Administrative	0.9%	75.0%	0.5%	1.2%	1.2%	0.2%	1.5%	0.5%	0.0%	1.9%	2.4%
Arts and Design	0.3%	1.3%	75.7%	2.0%	0.5%	1.0%	0.0%	4.6%	0.5%	0.5%	1.3%
Business Development	0.2%	2.0%	0.4%	78.6%	1.7%	0.0%	1.0%	1.1%	0.1%	0.9%	1.3%
Community and Social Services	0.0%	4.5%	0.0%	6.8%	65.0%	0.5%	1.4%	1.4%	0.0%	0.9%	4.5%
Consulting	3.2%	0.5%	0.2%	1.4%	0.2%	85.9%	0.2%	1.1%	0.2%	1.0%	0.8%
Education	0.9%	0.4%	0.7%	0.7%	0.7%	0.4%	87.3%	0.2%	0.0%	0.7%	2.6%
Engineering	0.0%	0.7%	0.8%	0.9%	0.3%	0.3%	0.3%	87.9%	0.2%	0.5%	0.5%
Entrepreneurship	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.9%	0.0%	94.9%	0.0%	1.7%
Finance	3.4%	1.5%	0.0%	1.3%	0.5%	0.5%	0.2%	0.3%	0.2%	85.9%	0.2%
Healthcare Services	0.0%	0.5%	0.2%	1.1%	0.6%	0.6%	1.5%	0.9%	0.0%	0.3%	87.2%
Human Resources	0.7%	2.0%	0.2%	3.8%	1.1%	2.2%	1.1%	0.0%	0.7%	0.0%	0.4%
Information Technology	0.6%	0.6%	0.6%	3.2%	0.0%	3.2%	0.3%	7.1%	1.0%	0.3%	0.0%
Legal	0.3%	1.6%	0.0%	1.0%	1.0%	0.3%	1.8%	0.5%	0.0%	2.6%	0.5%
Marketing	0.0%	0.0%	1.1%	0.4%	0.4%	0.0%	0.0%	0.2%	0.0%	0.0%	0.2%
Media and Communication	0.0%	1.0%	1.3%	2.9%	0.6%	0.3%	0.6%	1.0%	0.0%	0.3%	1.0%
Military and Protective Services	0.0%	0.0%	0.0%	3.6%	1.2%	0.0%	1.2%	2.4%	0.0%	6.0%	0.0%
Operations	0.6%	1.0%	0.8%	1.7%	1.1%	0.0%	0.4%	1.0%	0.0%	0.2%	1.0%
Product Management	0.3%	0.3%	0.8%	0.5%	0.0%	0.0%	0.0%	0.5%	0.3%	0.5%	0.3%
Program and Project Management	0.2%	0.4%	0.2%	1.8%	0.2%	0.0%	1.0%	2.1%	0.2%	0.0%	0.8%
Purchasing	0.0%	0.8%	0.0%	1.6%	0.8%	0.0%	0.0%	0.8%	0.0%	0.8%	0.0%

FIRST STEP: GET BETTER DATA



We share our models with devs and other teams

LinkedIn title classifier

Test the model

Title to classify:

account executive and marketing specialist

	Functions	Categories	Confidence
0	Inside Sales	Sales	1.0000
1	Sales	Sales	1.0000
2	Marketing	Marketing	1.0000

Source: custom dict

Naive classifier (for research only)



 [Submit your feedback](#)

Streamlit web app to share classifier with other teams

Request body required

Example Value | Schema

```
{
  "title": "Sales and Office Support"
}
```

Responses

Code Description

200

Successful Response

Media type

application/json

Controls Accept header.

Example Value | Schema

```
{
  "labels": {
    "sales": {
      "category": "master_sales",
      "score": 1
    },
    "office_operations": {
      "category": "master_operations",
      "score": 0.821090579032898
    }
  },
  "source": "custom dict + model"
}
```

Documentation for the API of the classifier

Results after deployment by devs into the database and integration in the searcher

Departments & Job Functions

Q Search departments...

▼ ☐ Sales

☐ Inside Sales

☐ Sales

☐ Business Development

☐ Revenue Operations

☐ Sales Engineering

☐ Sales Operations

☐ Partnerships

☐ Customer Success

☐ Channel Sales

☐ Account Management

☐ Sales Training

☐ Field / Outside Sales

☐ Sales Enablement

☐ Customer Retention Development

> ☐ Education

> ☐ Finance

- Millions of titles classified
- 13 high level categories / departments
- 196 subcategories
- Multiple categories per title possible

4.2. Company recommendation system

How can we help sales teams find their ideal customers?

- The Amplemarket platform
- A sales lead
- Lead qualification is manual
- Lots of time spent qualifying
- How can we support this process?

Company Name	Description	Potential Customer?
Jeronimo Martins	We have supermar..	✓
Facebook	A social media..	✗
Galp	Largest gas prov..	✓
EDP	Energias de Port..	✓
Google	At Google, we're..	✗
...

Lead qualification process
4 people every day

What the hell are embeddings?

1

Word Embeddings

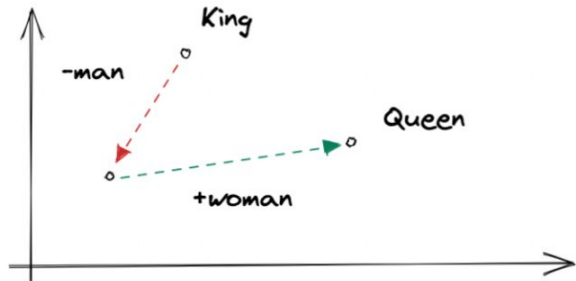
"King" → [0.67, -0.23, ...]

"Queen" → [-1.36, 0.29, ...]

"Woman" → [-2.67, 0.83, ...]

"Man" → [0.45, 0.91, ...]

2



3

Sentence Embeddings

"I" → [← N →]

"like" → [← N →]

"machine" → [← N →]

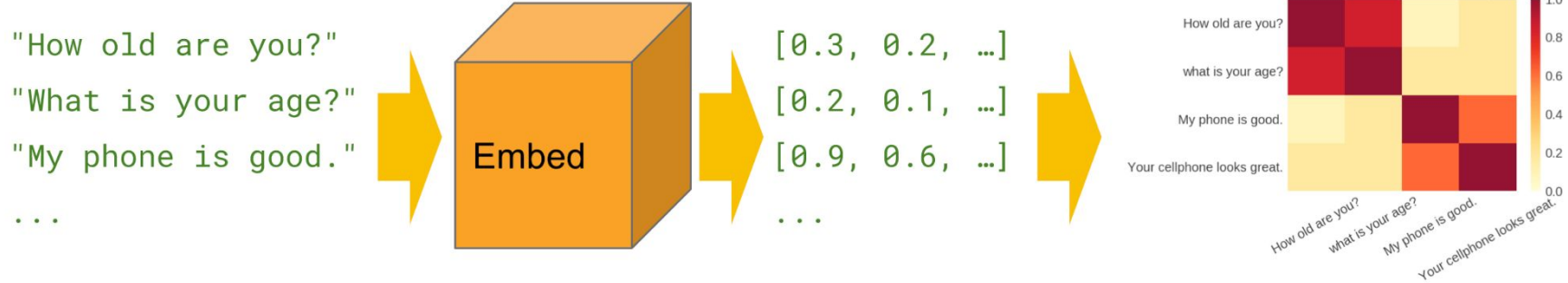
"learning" → [← N →]

Average

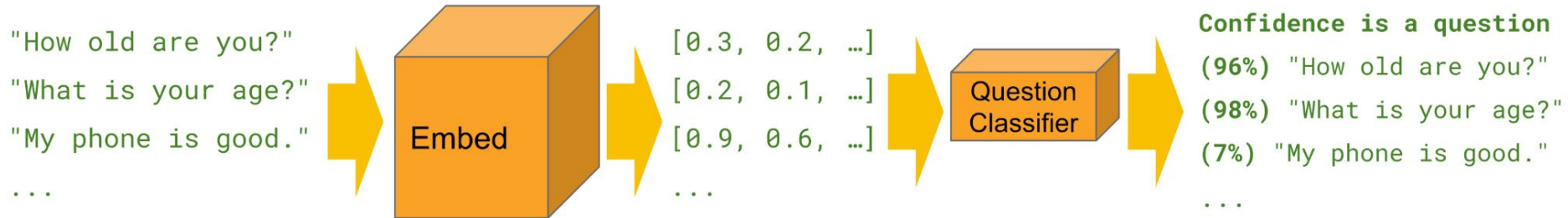
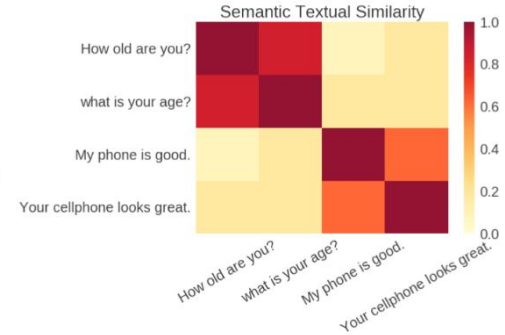
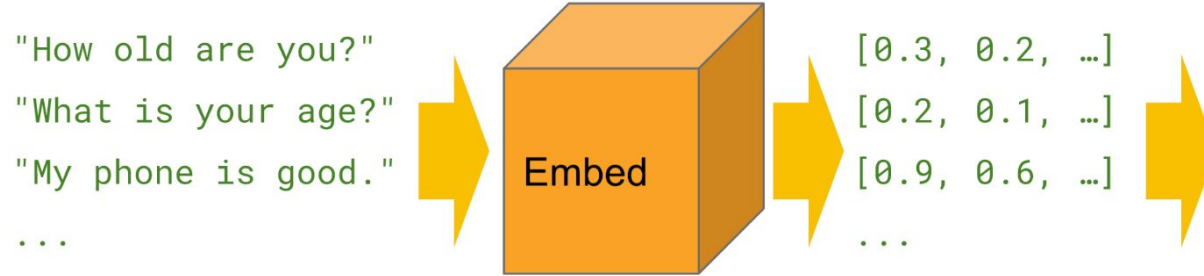
4

"I like machine learning" → [← N →]

Sentence embeddings have a wide range of applications in Machine Learning



Sentence embeddings have a wide range of applications in Machine Learning



5.Closing out

If you're interested in solving hard problems, reach out to us

- 100% remote and distributed
- From MIT, IST, DTU, etc
- Smart people all around
- amplemarket.com/carrers
- Or just come talk to us



Thank you!



Duarte O. Carmo

ML Engineer

duarte@amplemarket.com



Olavo Bacelar

ML Engineer

olavo@amplemarket.com

amplemarket

www.amplemarket.com

