

**Instituto de  
Computação**

UNIVERSIDADE ESTADUAL DE CAMPINAS



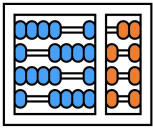
# Capacitação profissional em tecnologias de Inteligência Artificial

## **Machine Learning Overview**

**Prof. Edson Borin**

<https://www.ic.unicamp.br/~edson>

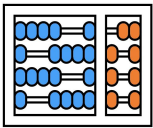
Institute of Computing - UNICAMP



# Machine Learning Overview



## ML Definition



# ML Definition: What is machine learning?

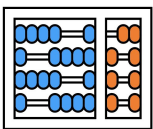


“Machine Learning is the **science (and art)** of programming computers so they can **learn from data**”.

—Aurélien Géron, 2019

“[Machine Learning is the] field of study that **gives computers the ability to learn without being explicitly programmed**”.

—Arthur Samuel, 1959

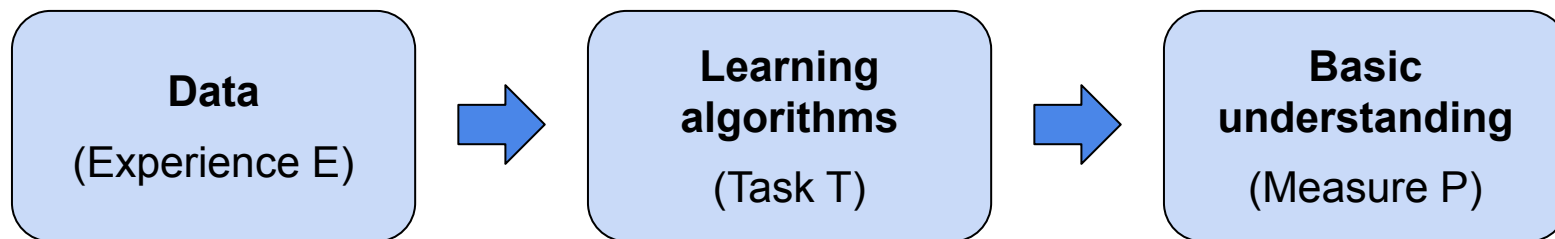


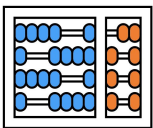
# ML Definition: What is machine learning?



“A computer program is said to learn from experience  $E$  with respect to some task  $T$  and some performance measure  $P$ , if its performance on  $T$ , as measured by  $P$ , improves with experience  $E$ ”.

—Tom Mitchell, 1997





# ML Definition: What is machine learning?



Example:



CC BY-SA (Christoph Scholz/Flickr)

**E: Regular and SPAM  
e-mail flagged by user**

**T: Classify whether  
email is SPAM or not**

**P: Classification  
accuracy**

**Data**  
(Experience E)



**Learning  
algorithms**  
(Task T)



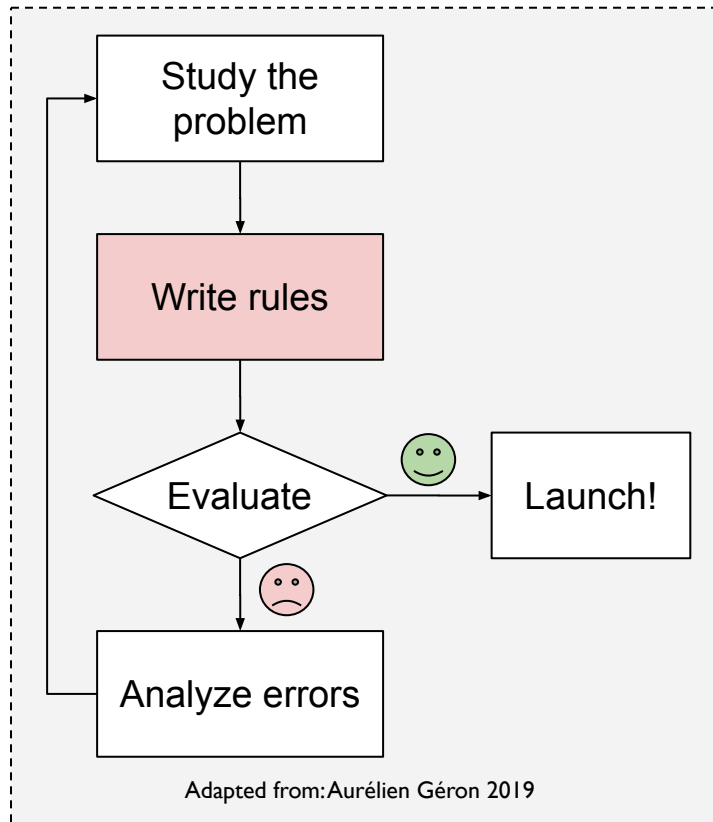
**Basic  
understanding**  
(Measure P)



# ML vs Rule-based algorithms

How does ML differ from traditional (rule-based) approach?

## Rule-based approach

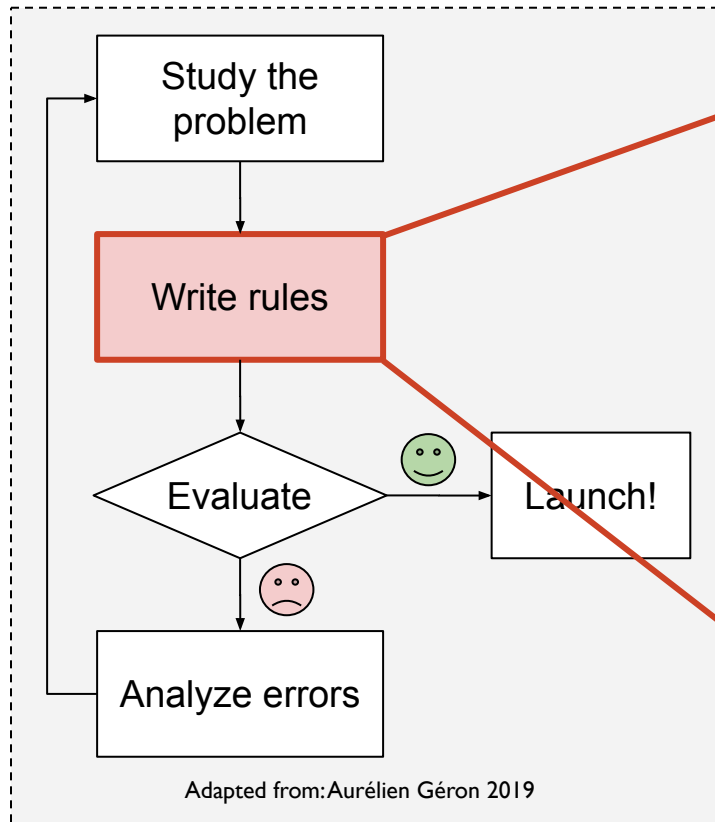




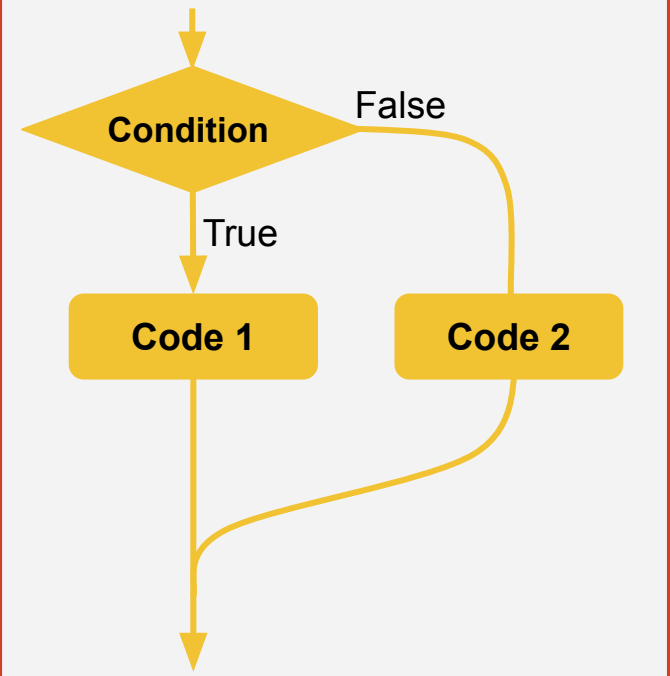
# ML vs Rule-based algorithms

How does ML differ from traditional (rule-based) approach?

## Rule-based approach



## Rule-based algorithm

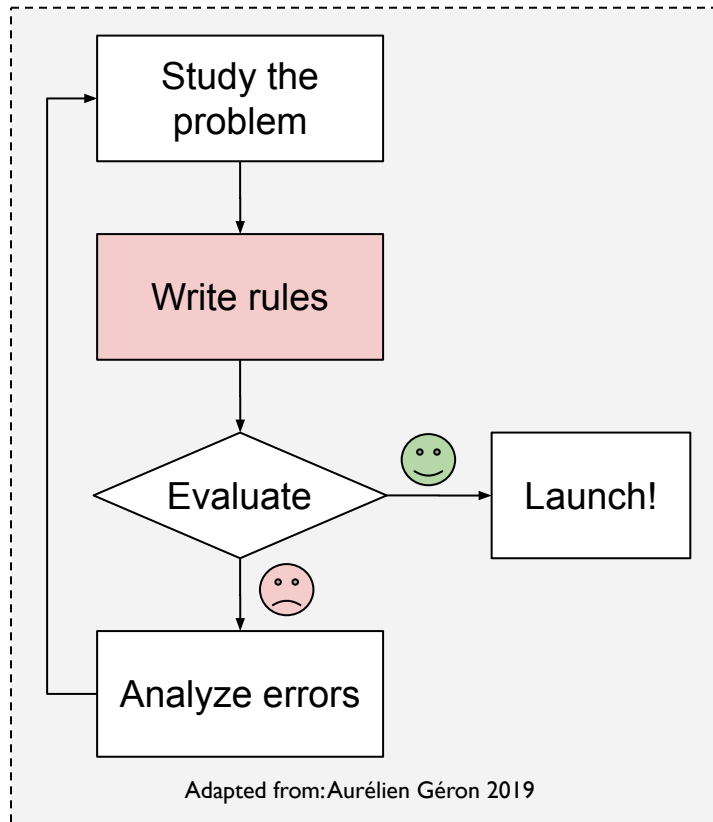




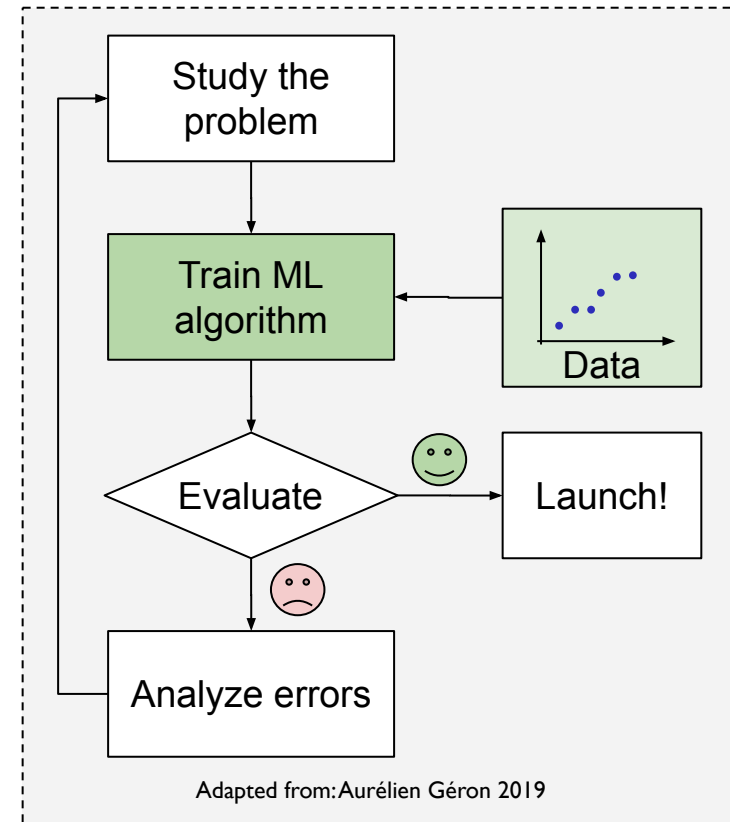
# ML vs Rule-based algorithms

How does ML differ from traditional (rule-based) approach?

## Rule-based approach



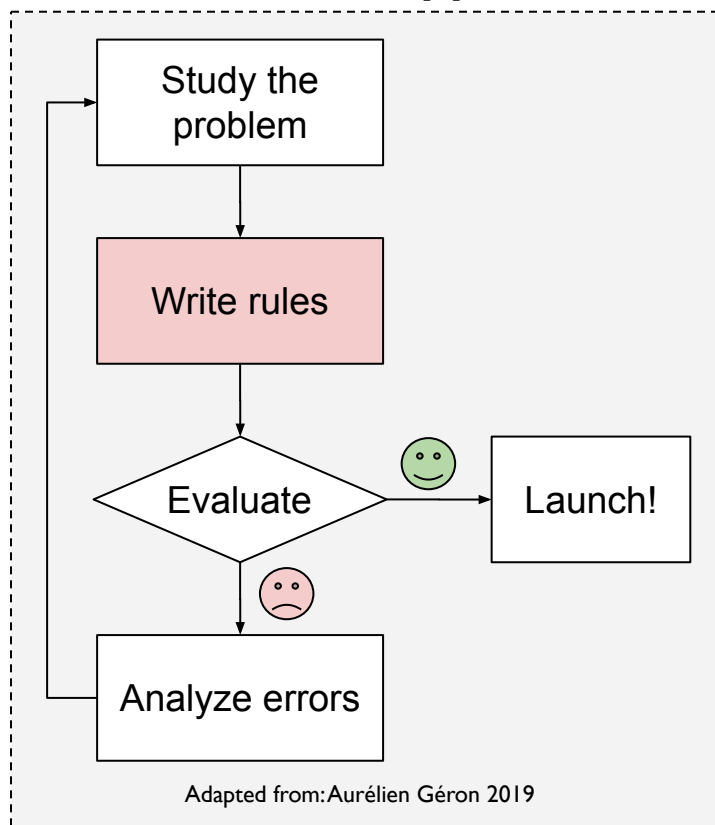
## ML approach





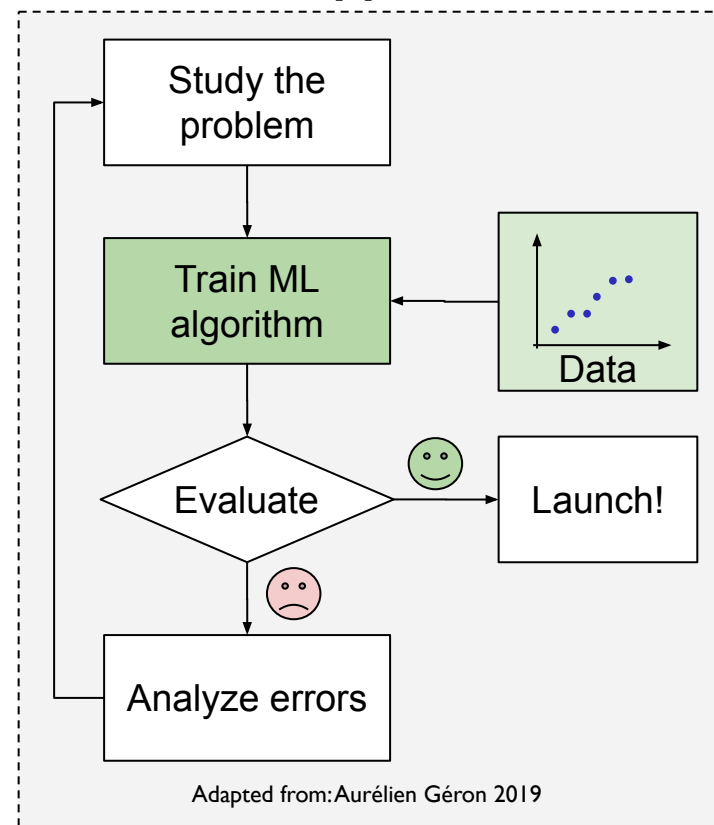
- Explicit programming is used to solve problems
- Rules are manually specified

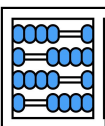
## Rule-based approach



- Samples are used for training
- Rules are automatically learned by machines

## ML approach





- Explicit programming is used to solve problems
- Rules are manually specified

ase

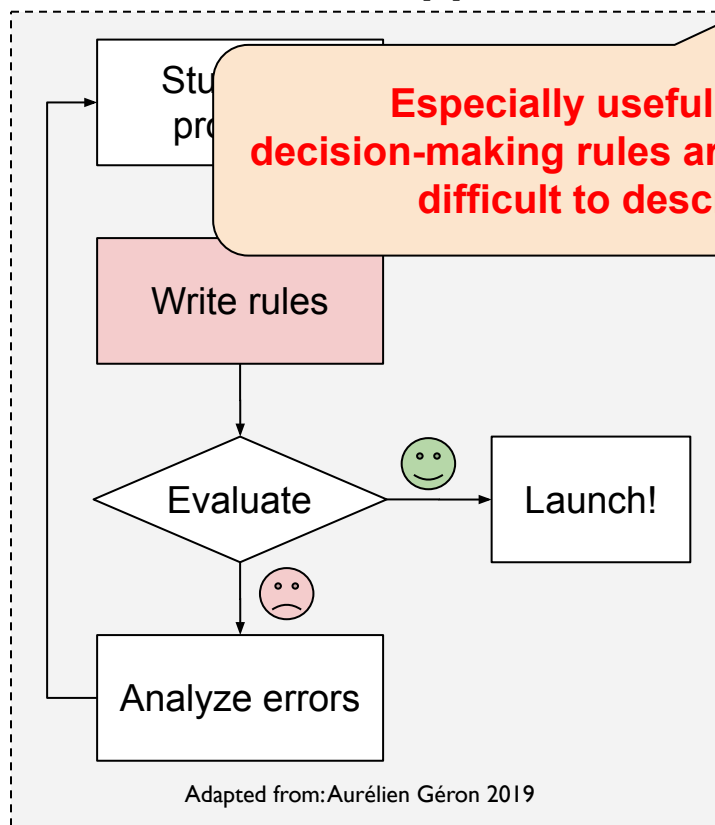
radit



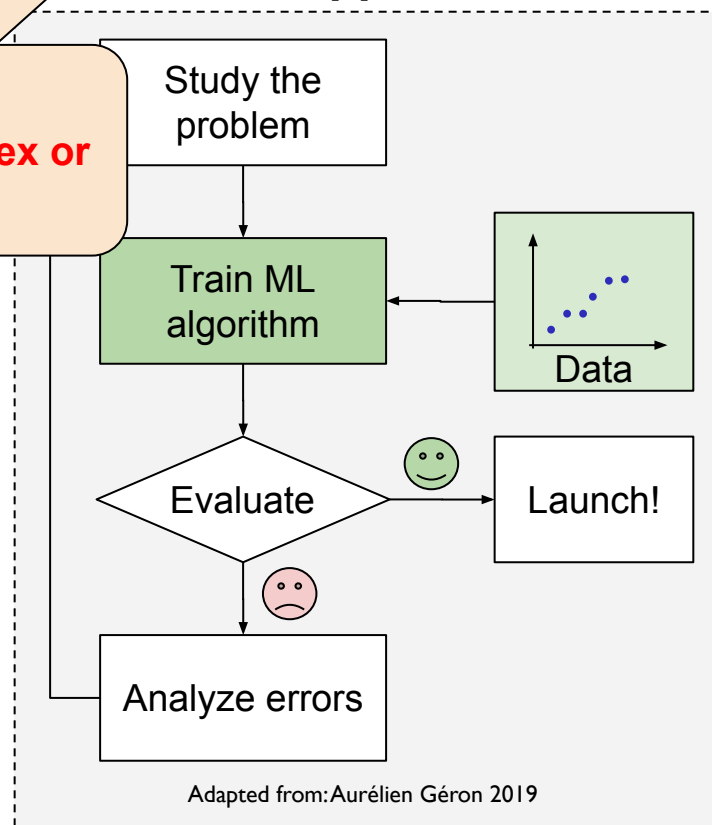
- Samples are used for training
- Rules are automatically learned by machines

h?

## Rule-based approach



## ML approach

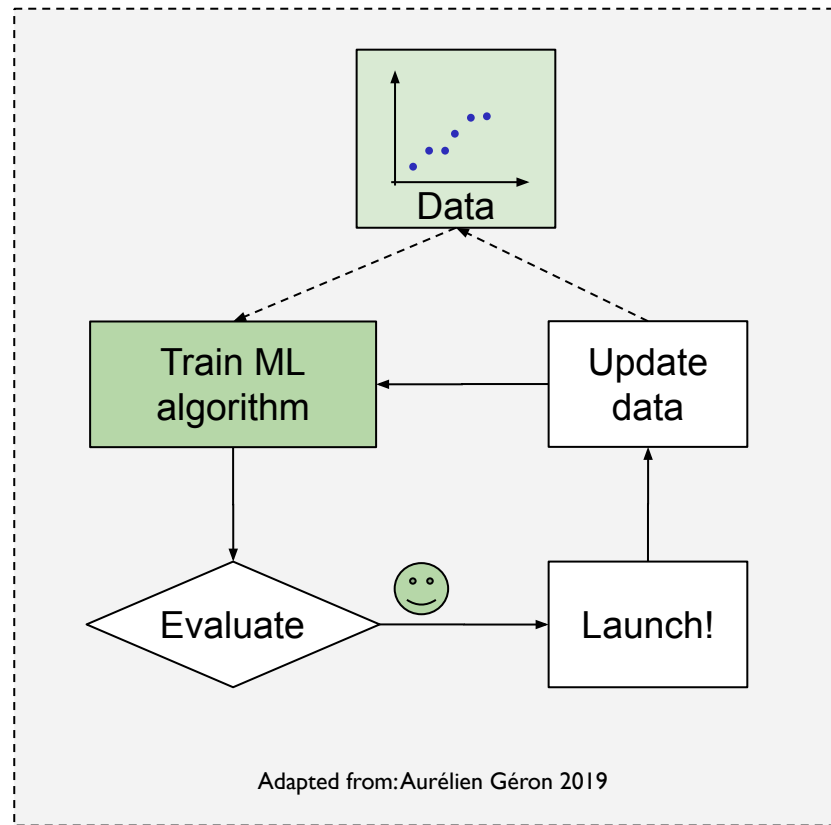


**Especially useful when  
decision-making rules are complex or  
difficult to describe!**



# ML vs Rule-based algorithms

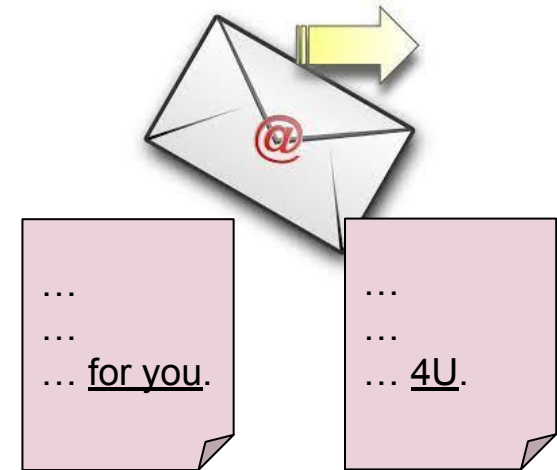
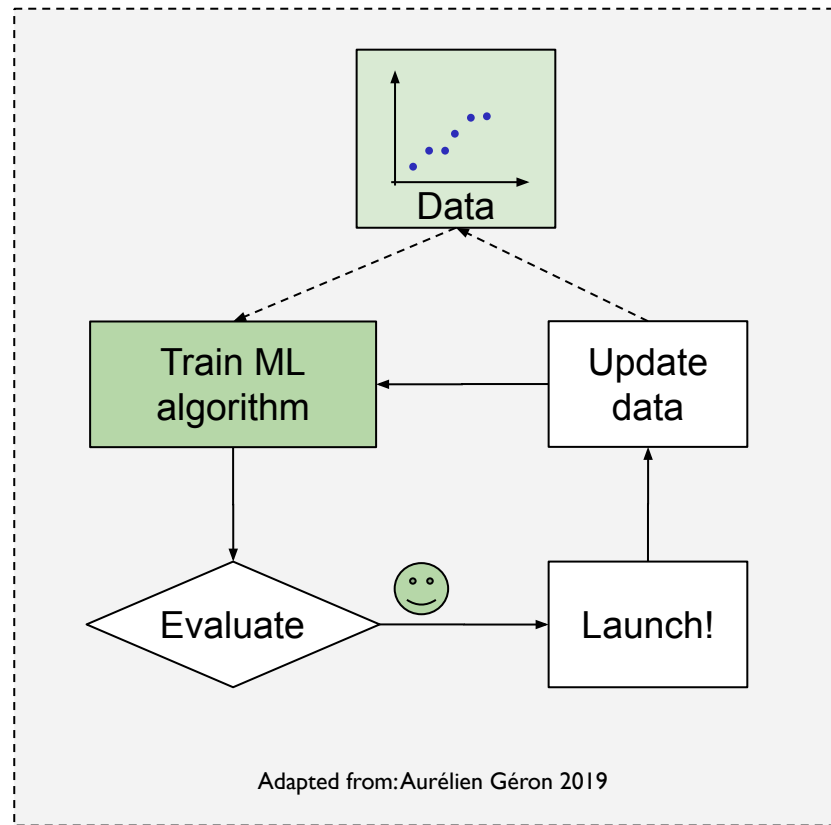
ML approach may also automatically adapt to changes!

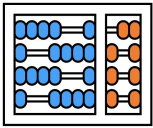




# ML vs Rule-based algorithms

ML approach may also automatically adapt to changes!



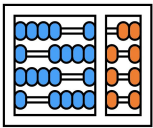


# Application scenarios of ML



## **ML is great for:**

- Problems for which existing solutions require a lot of hand-tuning or long lists of rules
  - One ML algorithm can often simplify code and perform better
- Complex problems for which there is no good solution at all using a traditional approach
  - The best ML techniques may find good solutions
- Fluctuating environments
  - A ML system can adapt to new data
- Getting insights about complex problems and large amounts of data



# Main Problems Solved by ML



- **Classification problems**

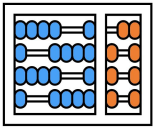
- A computer program (ML model) specify which of the  $k$  categories some input belongs to
- $f: R^n \rightarrow (1, 2, \dots, k)$
- Ex: Classify pictures

- **Regression problems**

- A computer program (ML model) predicts the output for the given input
- $f: R^n \rightarrow R$
- Ex: Estimate insurance premium or house price

- **Clustering**

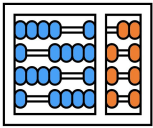
- A computer program divides the samples of a dataset into multiple categories according to internal similarity of the data
- Ex: Identify similar users based on their profile



# Machine Learning Overview



## ML Types



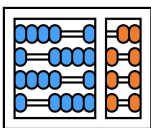
# ML Definition



## Agenda

- Supervised learning
- Unsupervised learning
- Semi-supervised learning
- Reinforcement learning
- Batch and Online learning
- Generalization approach



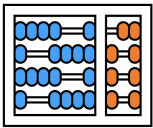


# ML Types



ML systems are usually classified according to:

- Whether or not they are trained with human supervision (Supervised, unsupervised, semi-supervised, and Reinforcement learning)
- Whether or not they can learn incrementally on the fly (online versus batch learning)
- Whether they work by simply comparing new data points to known data points or by detecting patterns in the training data and building a predictive model (instance-based versus model-based learning).

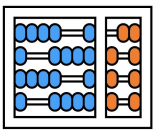


# ML Types



ML systems are usually classified according to:

- Whether or not they are trained with human supervision (Supervised, unsupervised, semi-supervised, and Reinforcement learning)
- Whether or not they can learn incrementally on the fly (online versus batch learning)
- Whether they work by simply comparing new data points to known data points or by detecting patterns in the training data and building a predictive model (instance-based versus model-based learning).



# ML Types



ML systems are usually classified according to:

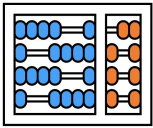
- Whether or not they are trained with human supervision (Supervised, unsupervised, semi-supervised, and Reinforcement learning)
- Whether or not they can learn incrementally on the fly (online versus batch learning)
- Whether they work by simply comparing new data points to known data points or by detecting patterns in the training data and building a predictive model (instance-based versus model-based learning).



# Machine Learning Overview

## ML Types

### Supervised learning

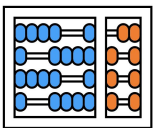


# ML Types: Supervised learning



In the **supervised learning**, the data used to train the model is labeled with the desired/expected solution.

- Used both for classification and regression tasks!



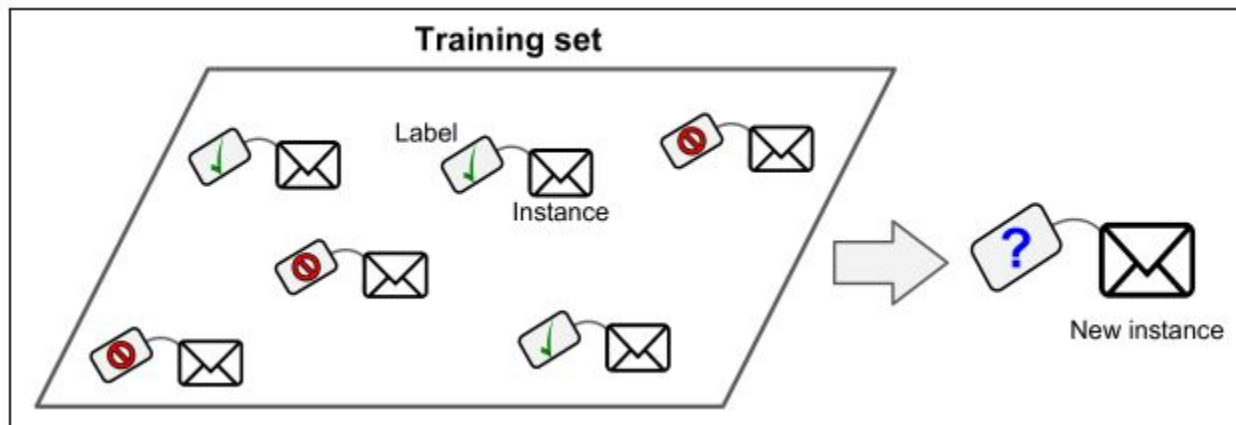
# ML Types: Supervised learning



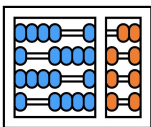
In the **supervised learning**, the data used to train the model is labeled with the desired/expected solution.

## Example of classification task

- Task: Spam classification
- Training data: Email messages labeled SPAM or NOT-SPAM



Source: Aurélien Géron 2019



# ML Types: Supervised learning



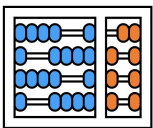
In the **supervised learning**, the data used to train the model is labeled with the desired/expected solution.

## Example of classification task

- Task: Handwritten digits recognition
- Training data: Images of handwritten digits labeled with their respective digit class (0, 1, 2, ..., or 9)



Sample images from  
MNIST dataset



# ML Types: Supervised learning



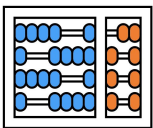
In the **supervised learning**, the data used to train the model is labeled with the desired/expected solution.

## Example of regression task

- Task: Predict price of a car
- Training data: Car attributes (model, mileage, age, etc.) labeled with their respective market price (Price)

| Model   | Mileage | Age | Price      |
|---------|---------|-----|------------|
| Ford Ka | 5,000   | 1   | 10,000 USD |
| Ford Ka | 40,000  | 2   | 7,000 USD  |
| Fusca   | 300,000 | 30  | 800 USD    |





# ML Types: Supervised learning



In the **supervised learning**, the model is labeled with the **desired solution**.

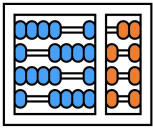
On regression tasks, the output is a **numeric value** (instead of a class identifier)

**Ex:** How much would a year old 20k miles Ford Ka cost?

## Example of regression task

- Task: Predict price of a car
- Training data: Car attributes (model, mileage, age, etc.) labeled with their respective market price (Price)

| Model   | Mileage | Age | Price      |
|---------|---------|-----|------------|
| Ford Ka | 5,000   | 1   | 10,000 USD |
| Ford Ka | 40,000  | 2   | 7,000 USD  |
| Fusca   | 300,000 | 30  | 800 USD    |



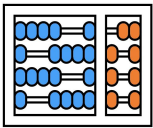
# ML Types: Supervised learning



In the **supervised learning**, the data used to train the model is labeled with the desired/expected solution.

## Common tasks that can be solved by supervised learning

- Classification
- Regression
- Anomaly Detection
- Synthesis and sampling
- Denoising
- ...



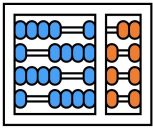
# ML Types: Supervised learning



In the **supervised learning**, the data used to train the model is labeled with the desired/expected solution.

Examples of supervised learning algorithms:

- k-Nearest Neighbors
- Linear Regression
- Logistic Regression
- Support Vector Machines (SVMs)
- Decision Trees and Random Forests
- Neural networks

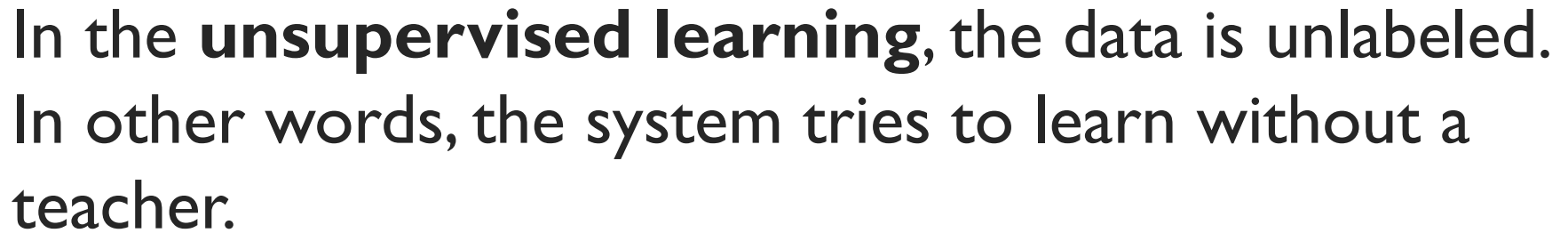


# Machine Learning Overview



## ML Types

### Unsupervised learning



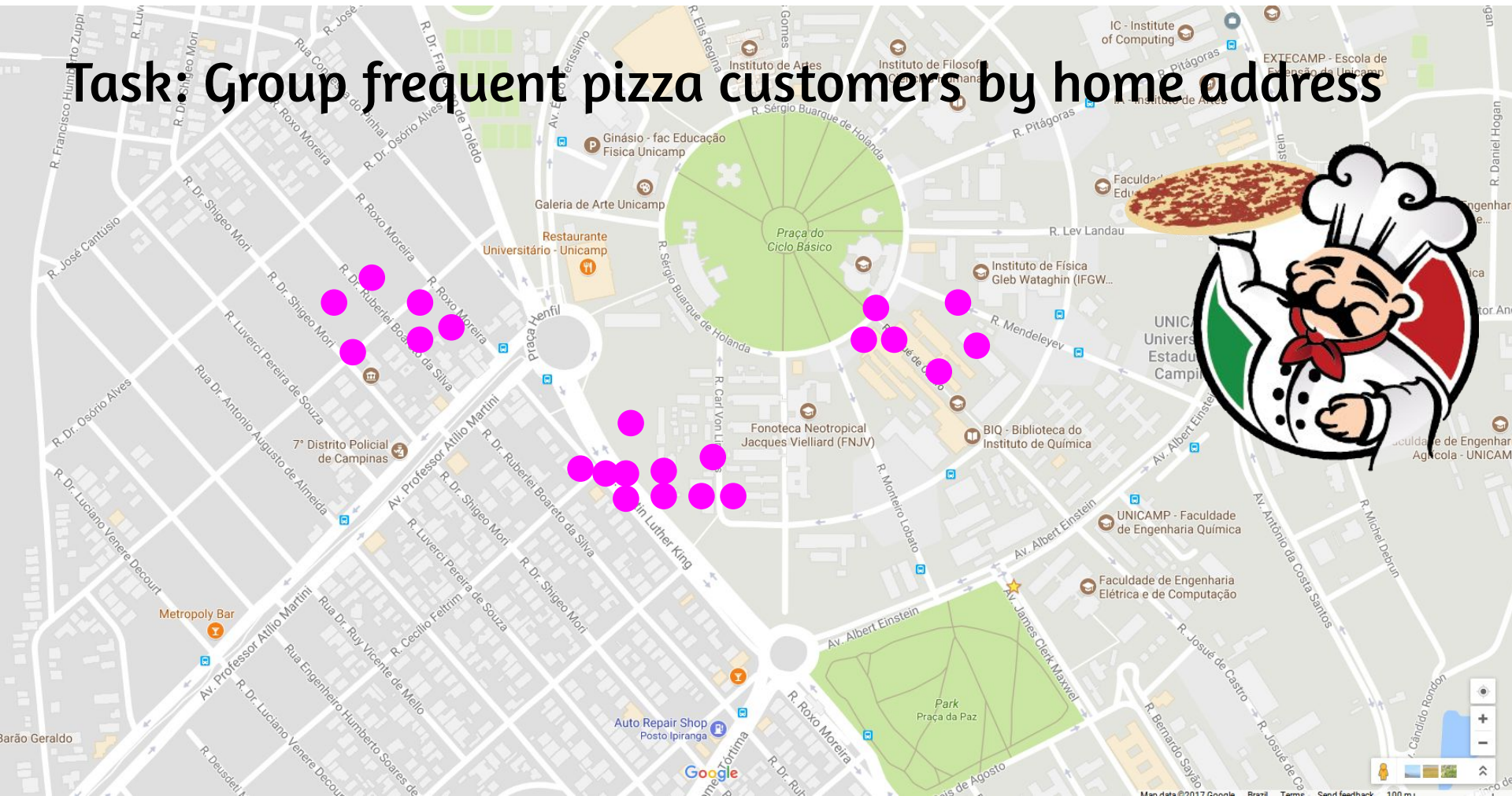
- Task: Group frequent pizza customers by home address
- Training data: Frequent consumers' **addresses**





# ML Types: Unsupervised learning

Task: Group frequent pizza customers by home address



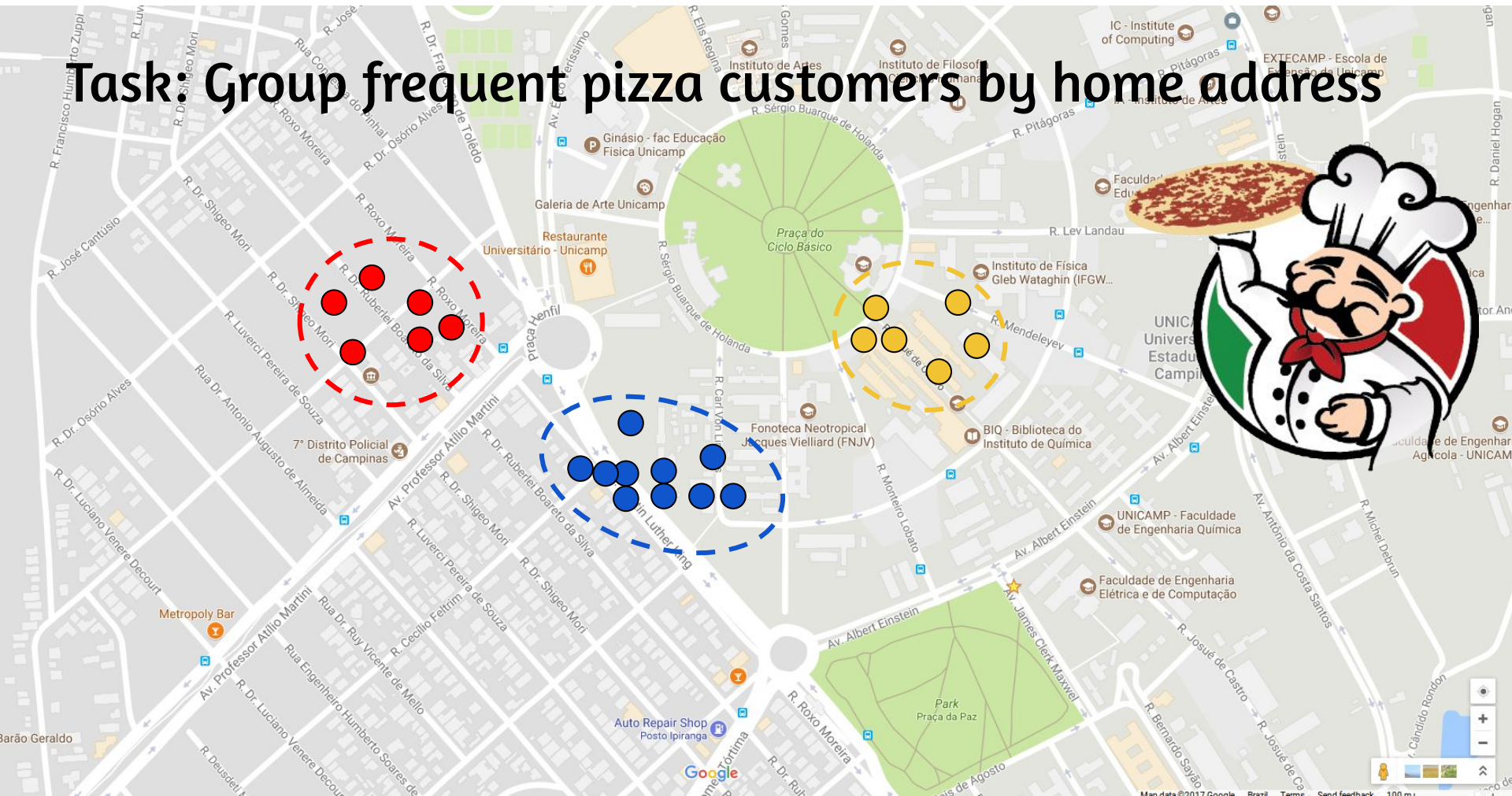
Adapted from: <https://www.youtube.com/watch?v=lpGxLWOIZy4> and Prof. Sandra Avila slides





# ML Types: Unsupervised learning

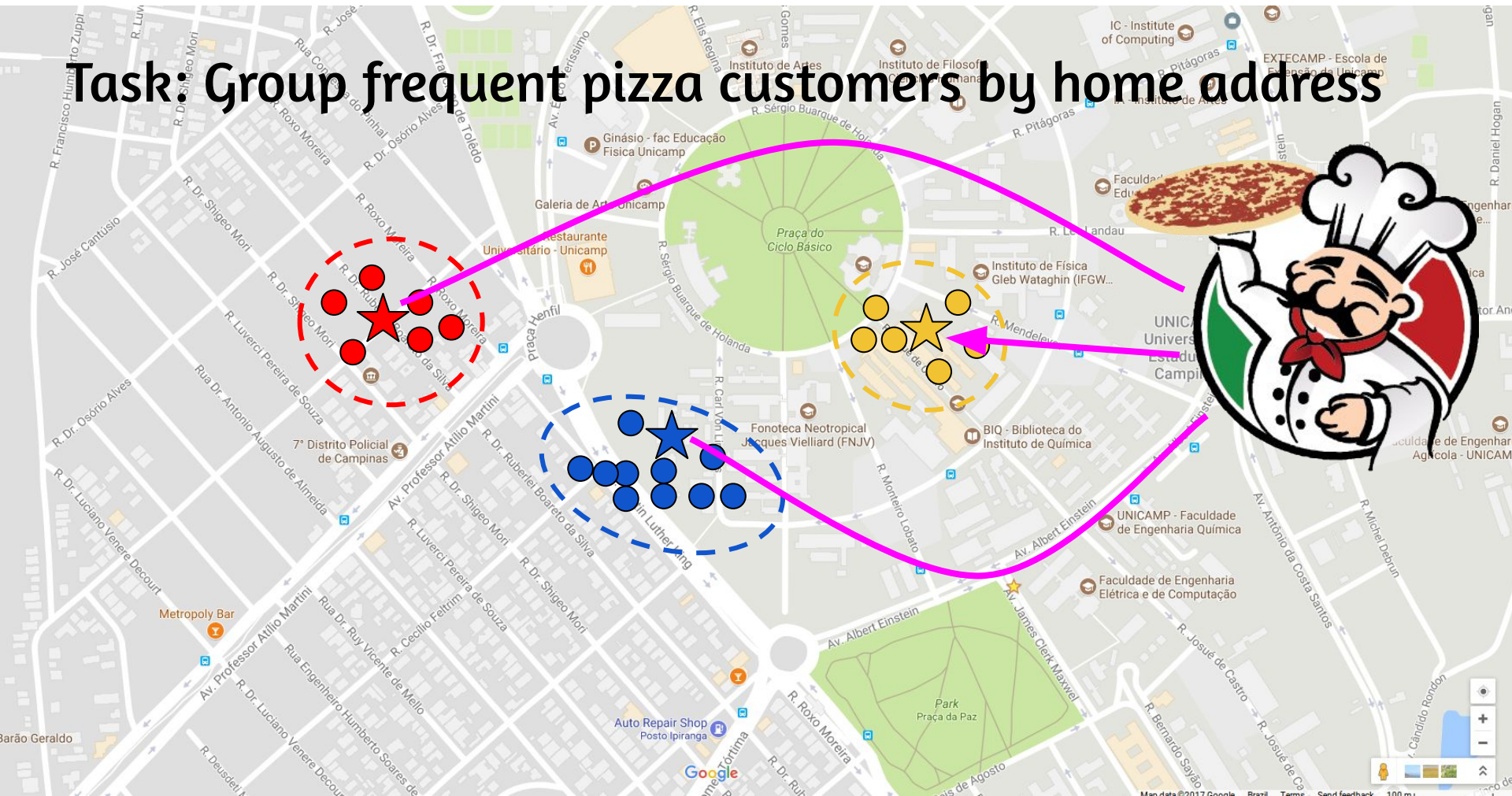
**Task: Group frequent pizza customers by home address**





# ML Types: Unsupervised learning

Task: Group frequent pizza customers by home address

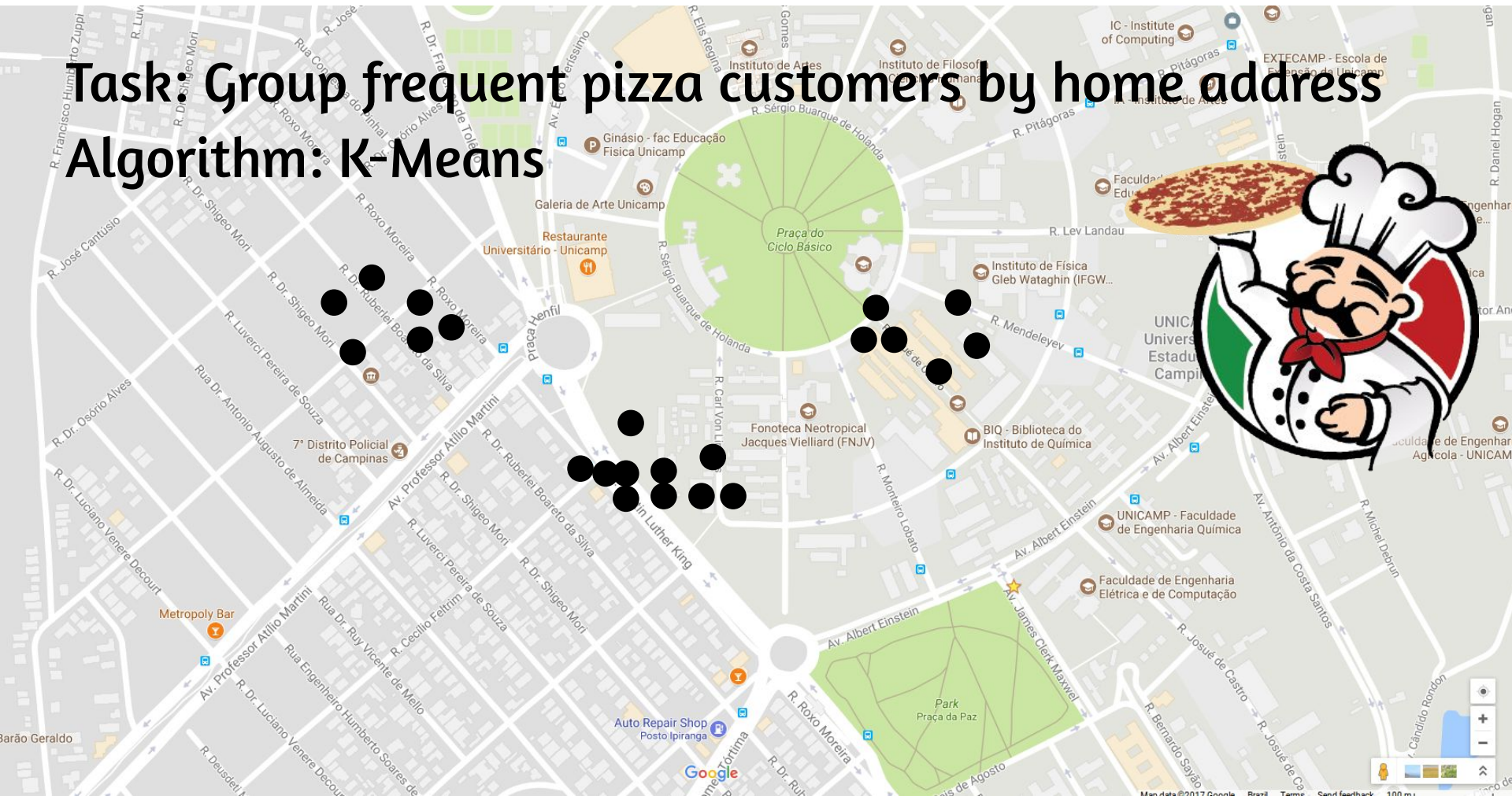






# ML Types: Unsupervised learning

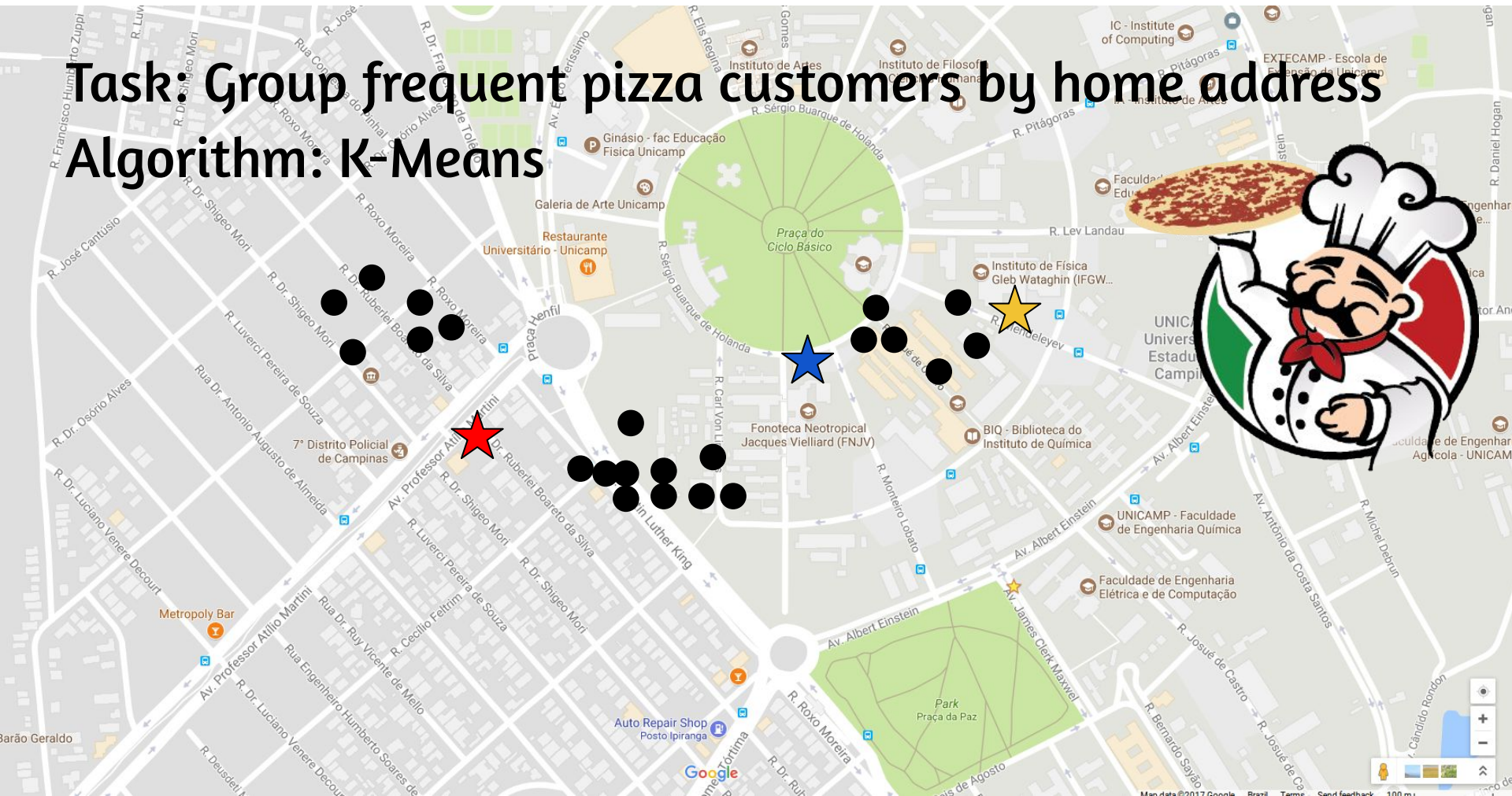
**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**





# ML Types: Unsupervised learning

**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**

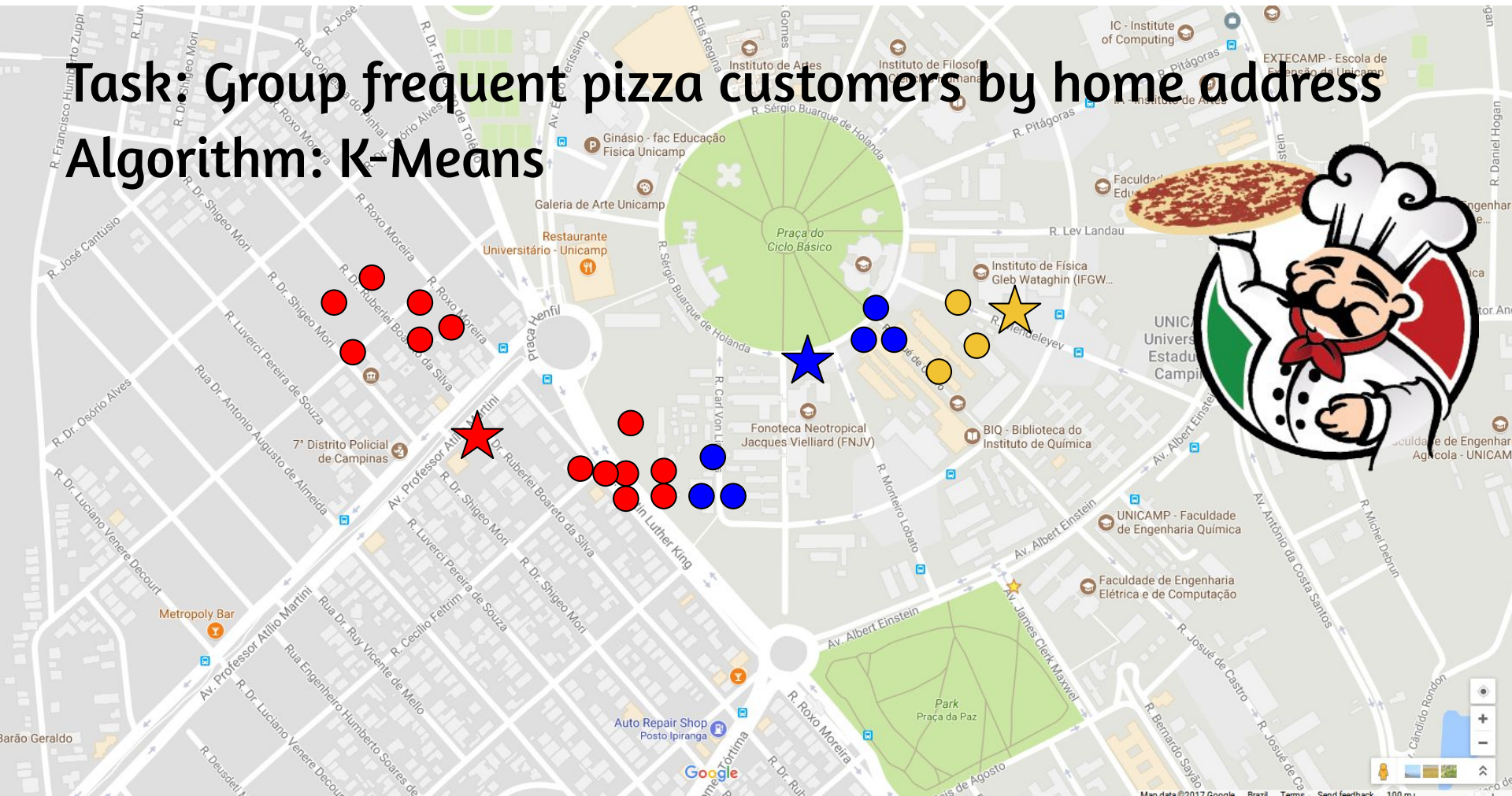






# ML Types: Unsupervised learning

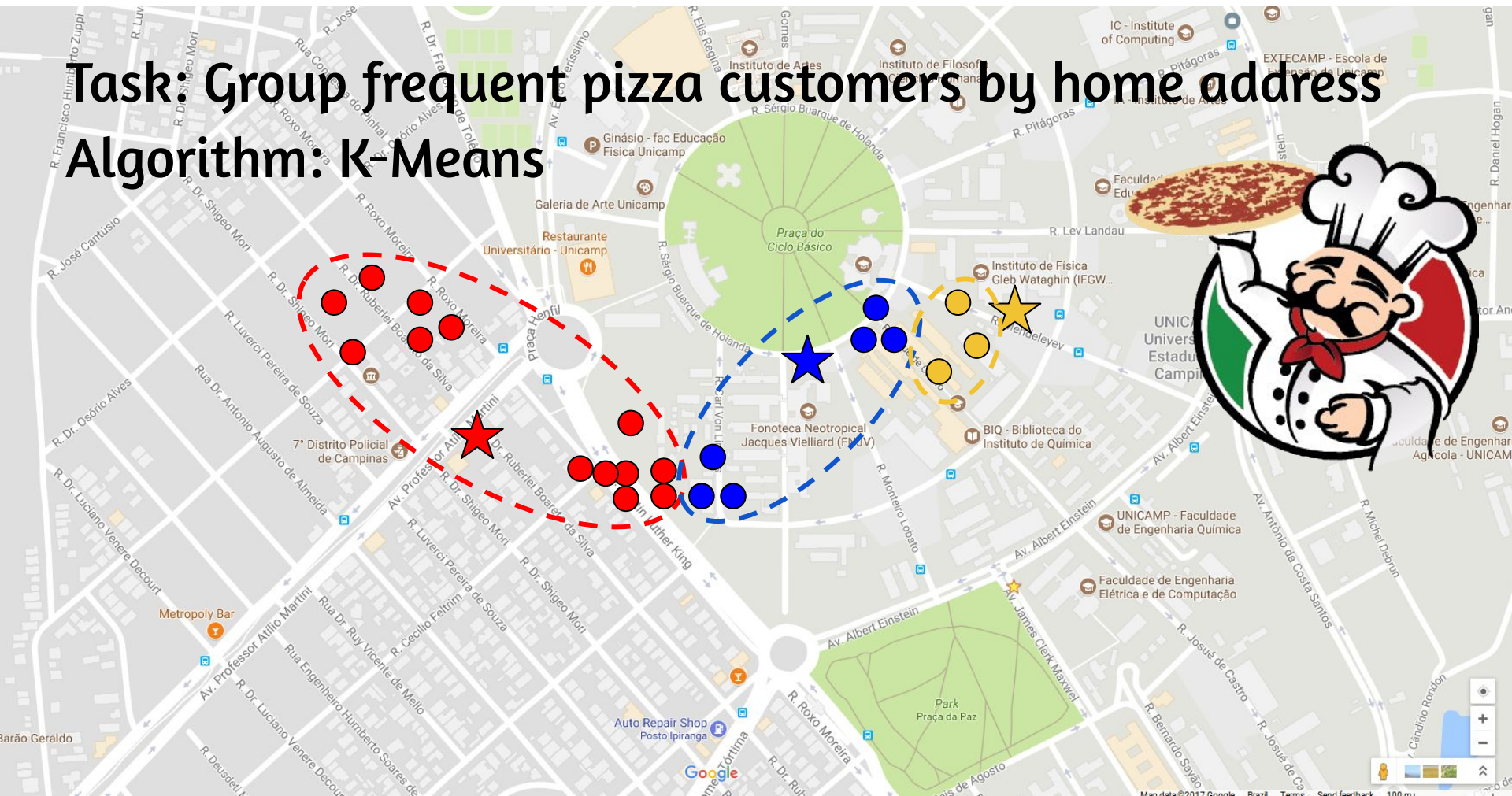
**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**





# ML Types: Unsupervised learning

**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**

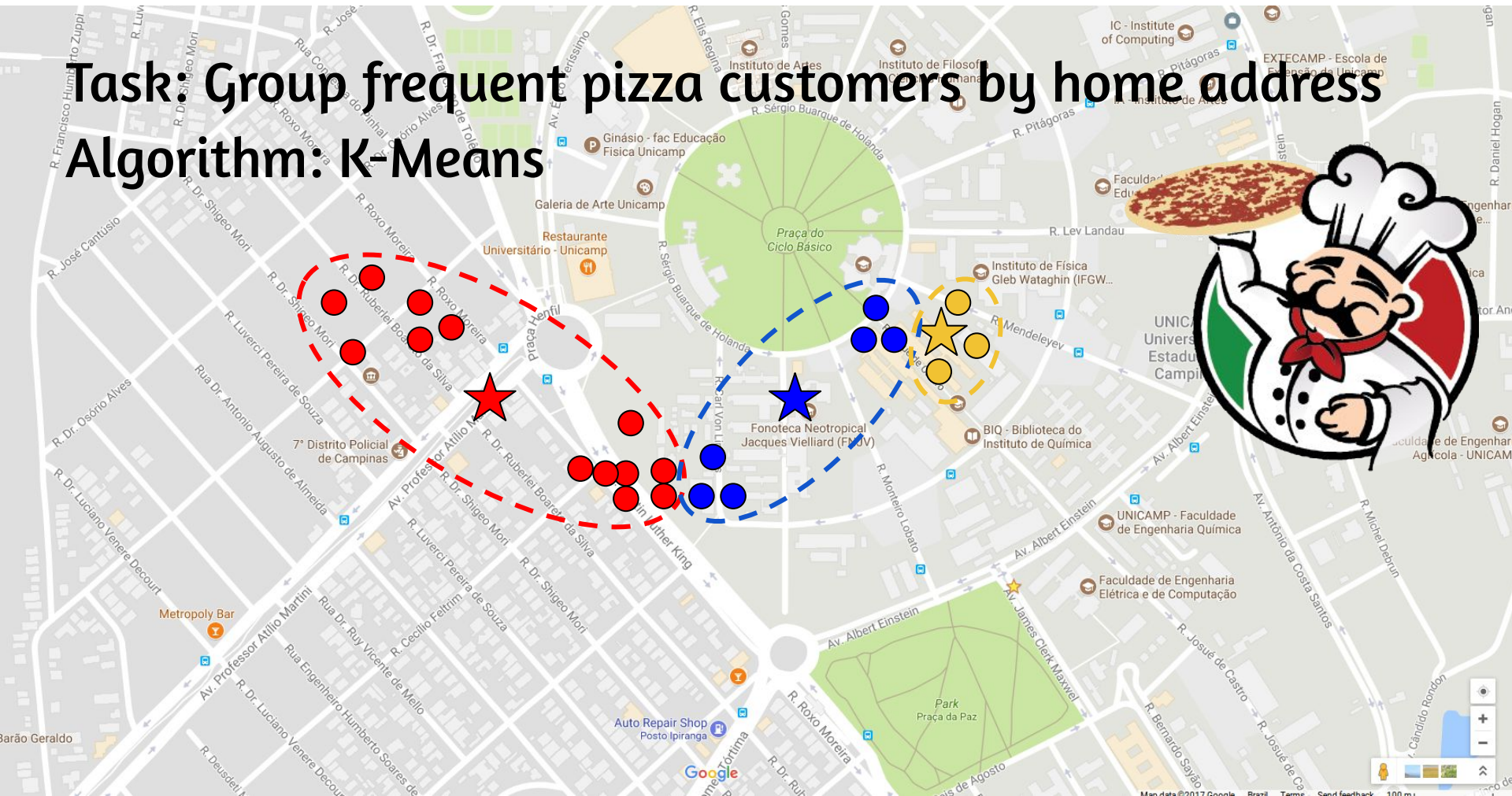






# ML Types: Unsupervised learning

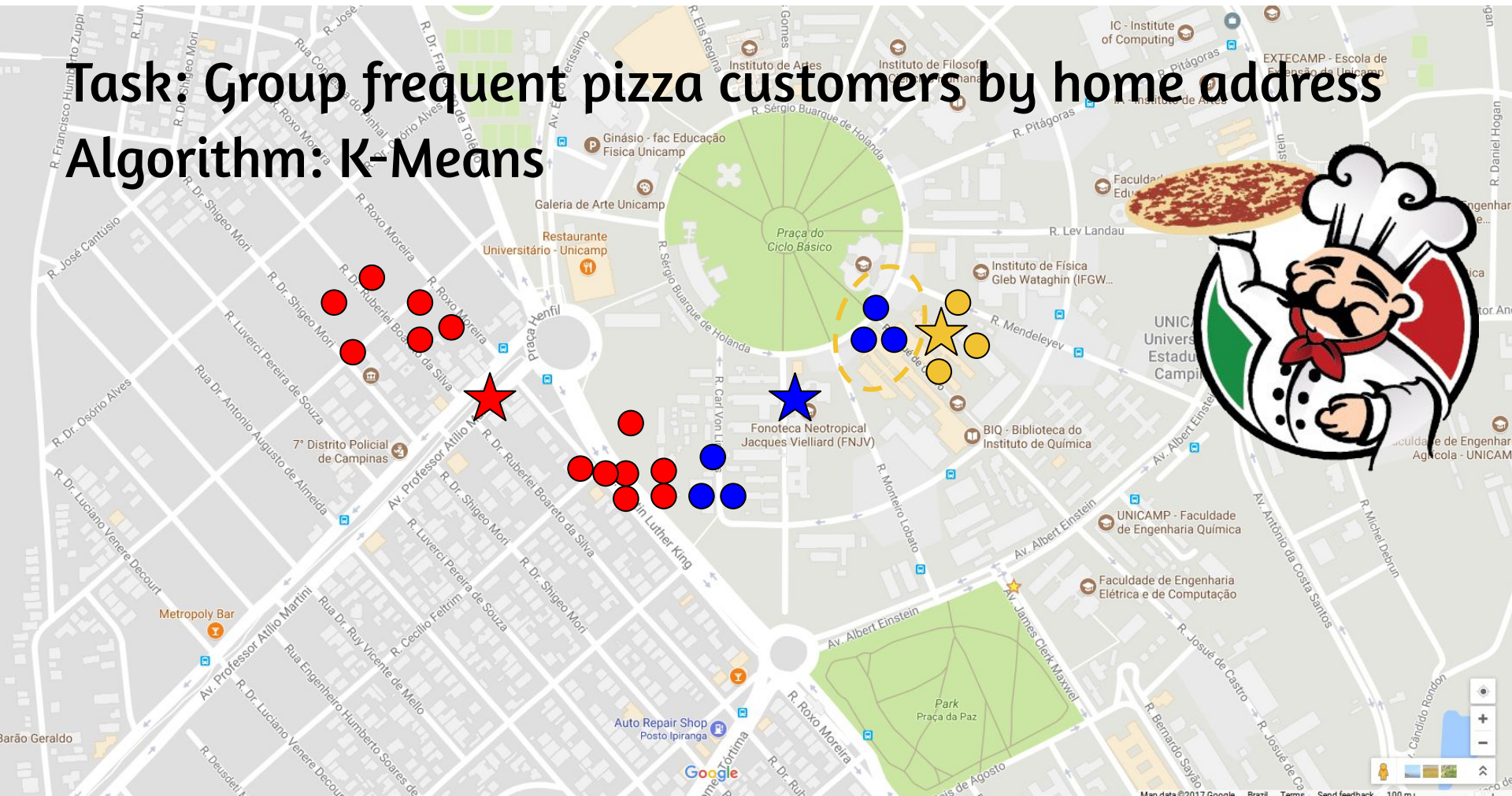
**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**





# ML Types: Unsupervised learning

**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**

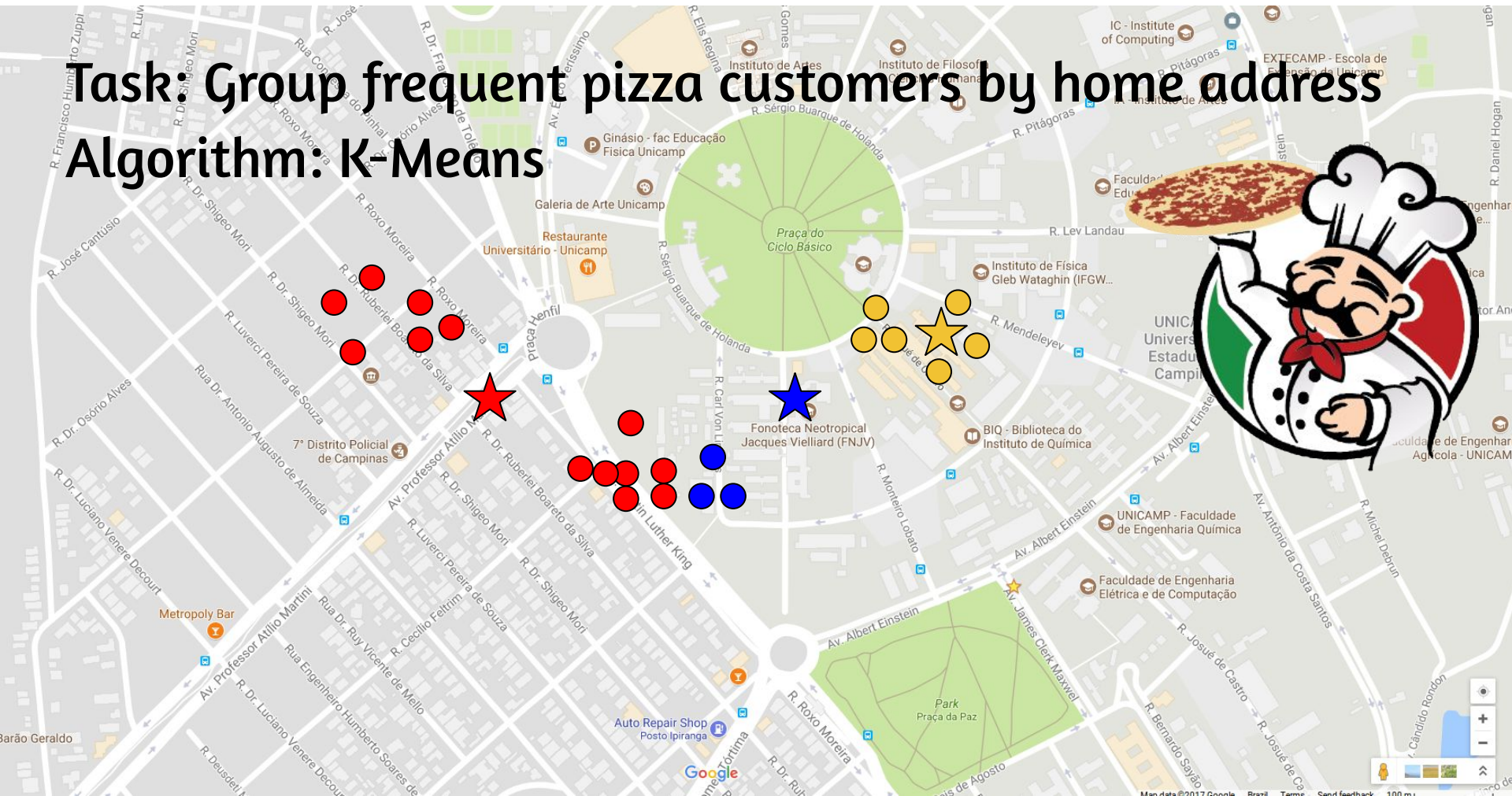






# ML Types: Unsupervised learning

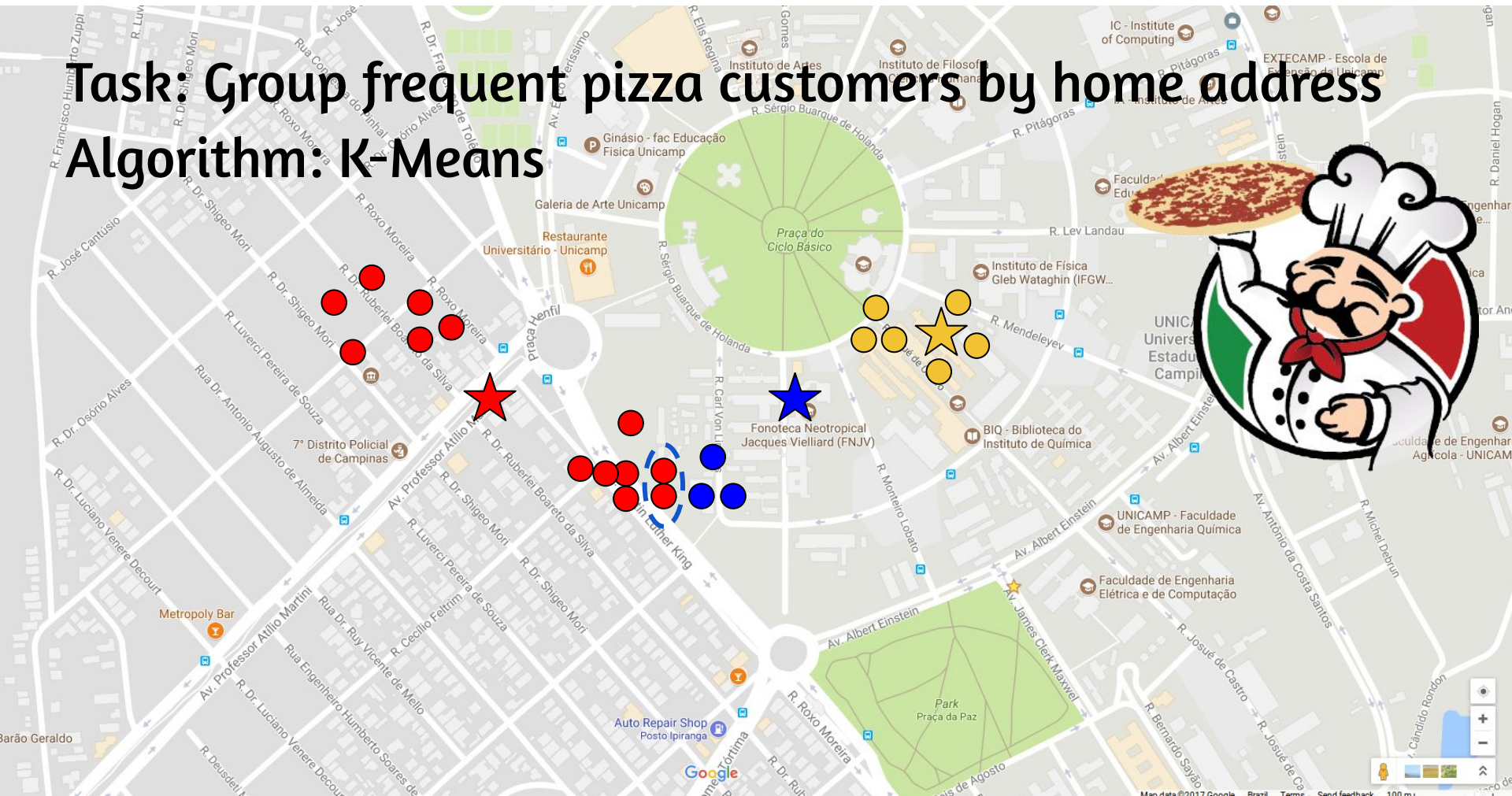
**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**





# ML Types: Unsupervised learning

**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**

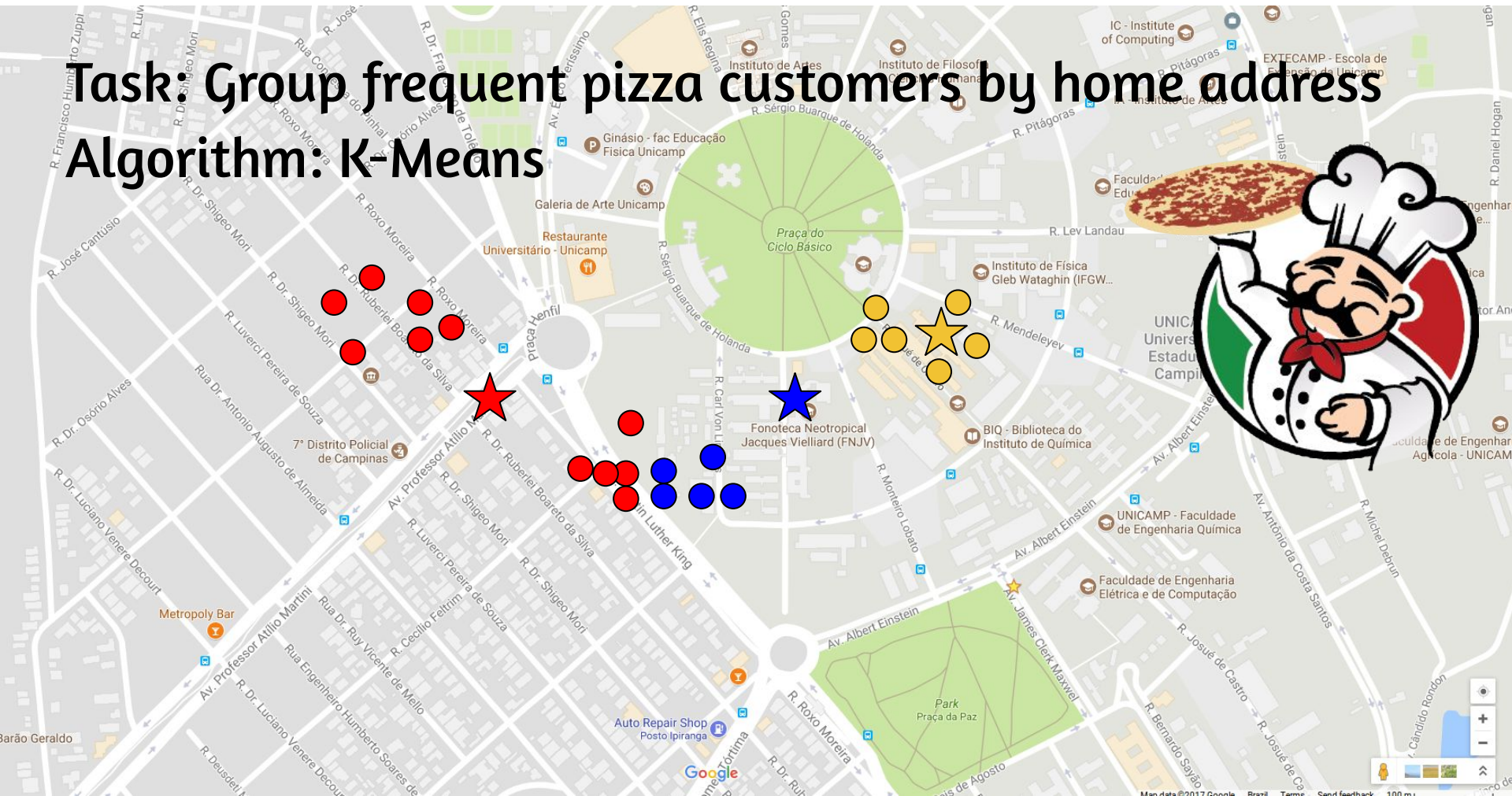






# ML Types: Unsupervised learning

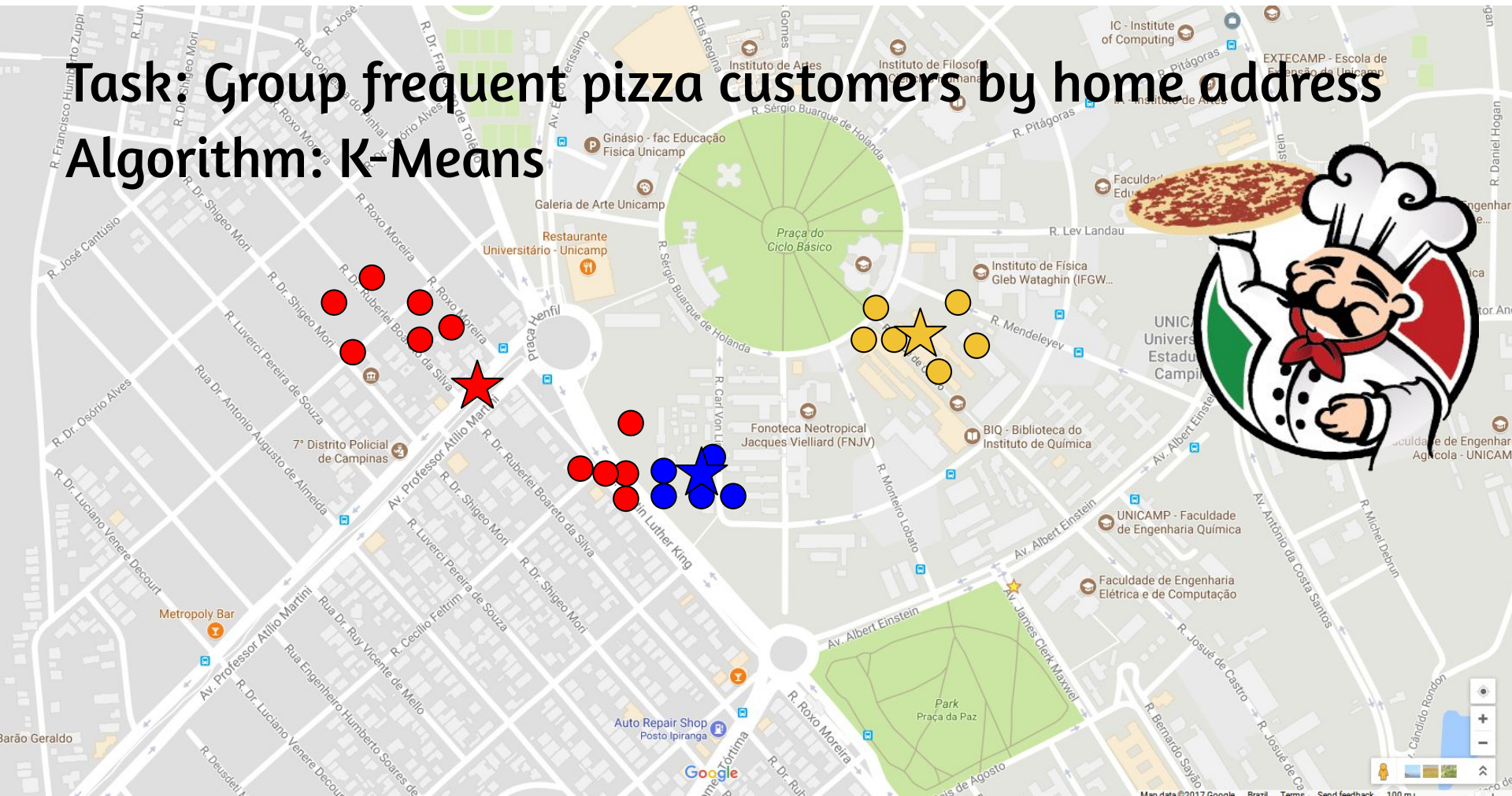
**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**





# ML Types: Unsupervised learning

**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**

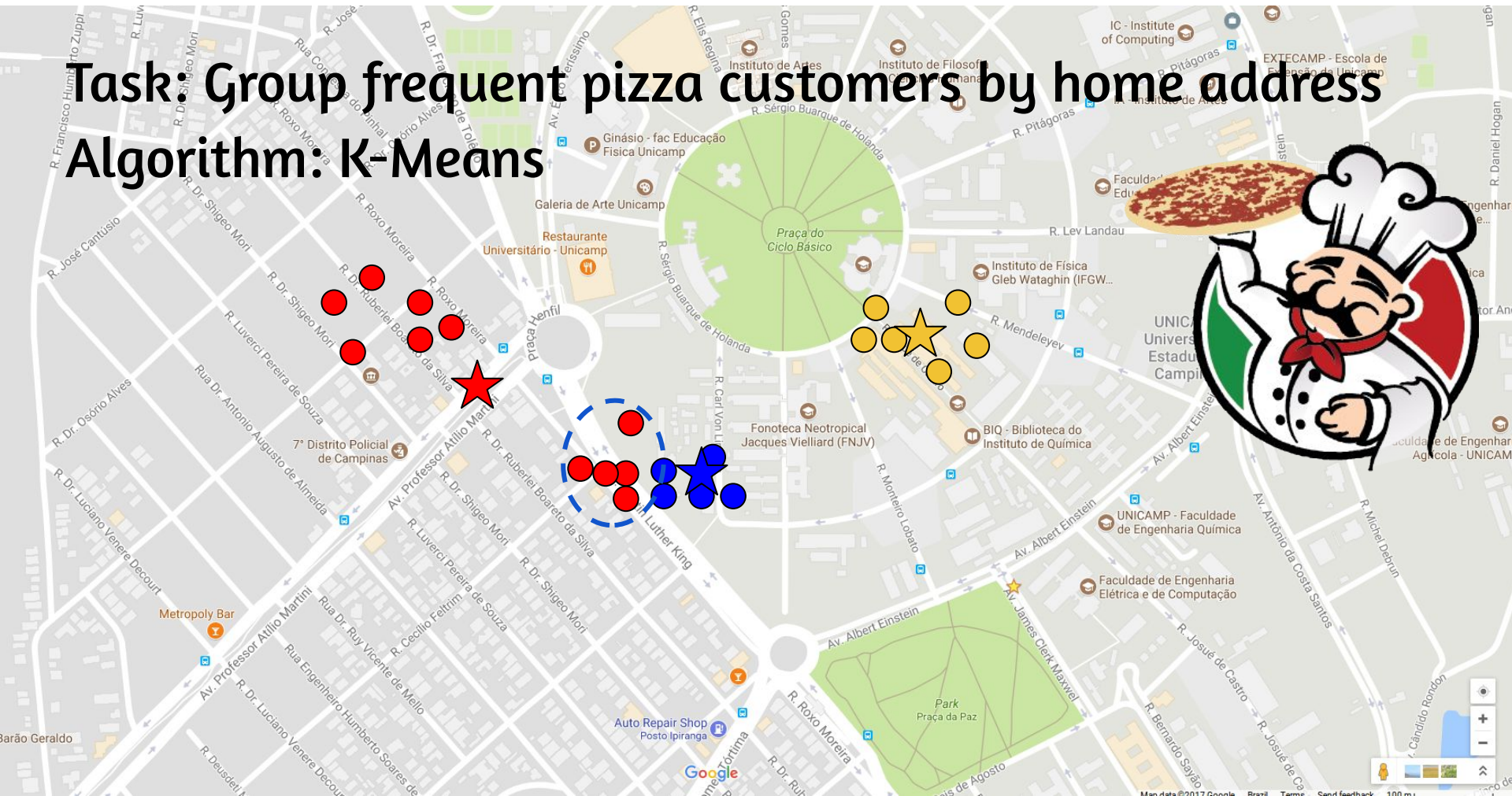






# ML Types: Unsupervised learning

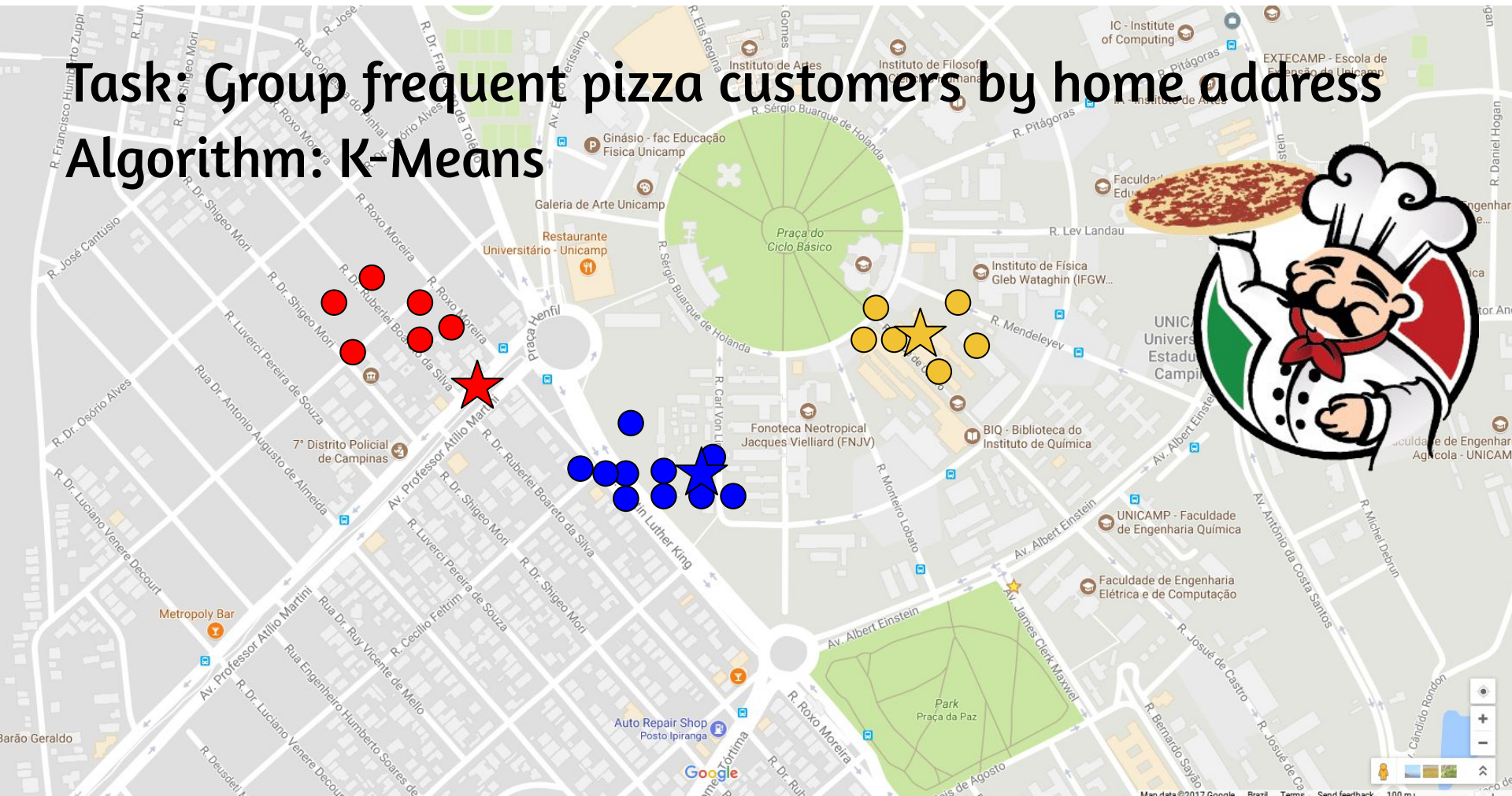
**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**





# ML Types: Unsupervised learning

**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**

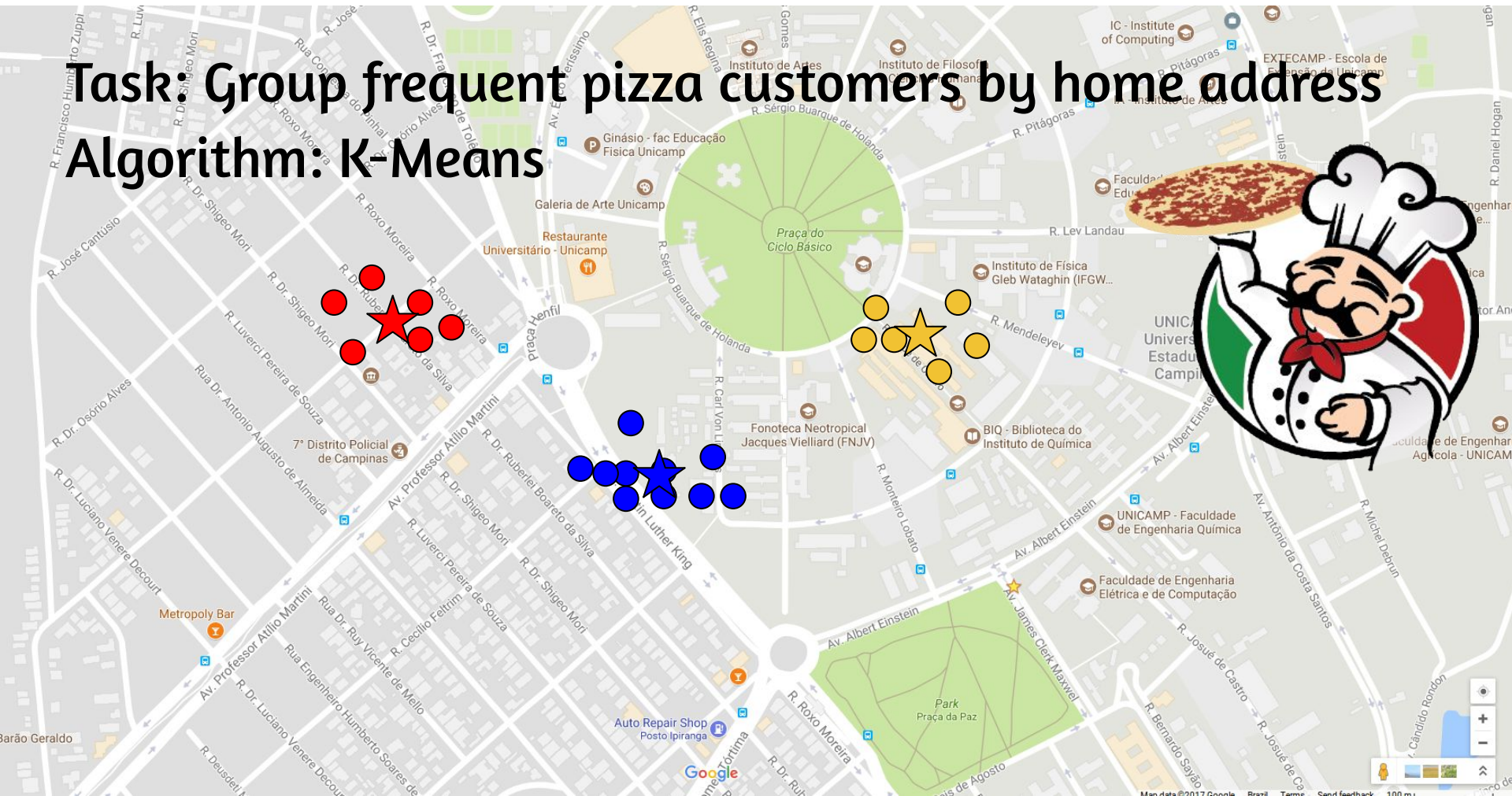


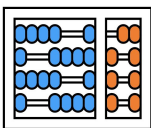




# ML Types: Unsupervised learning

**Task: Group frequent pizza customers by home address**  
**Algorithm: K-Means**





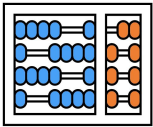
# ML Types: Unsupervised learning



In the **unsupervised learning**, the data is unlabeled. In other words, the system tries to learn without a teacher.

Examples of unsupervised learning algorithms:

- Clustering: K-Means, DBSCAN, Hierarchical Cluster Analysis (HCA)
- Visualization and dimensionality reduction: Principal Component Analysis (PCA), t-distributed Stochastic Neighbor Embedding (t-SNE), UMAP, etc.

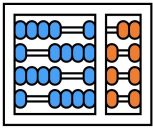


# Machine Learning Overview



## ML Types

### Semi-supervised learning



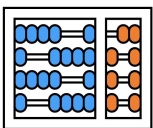
# ML Types: Semi-supervised learning



In the **semi-supervised learning**, only a subset of the dataset used to train the model is labeled with the desired/expected solution.

- Used both for classification and regression tasks!





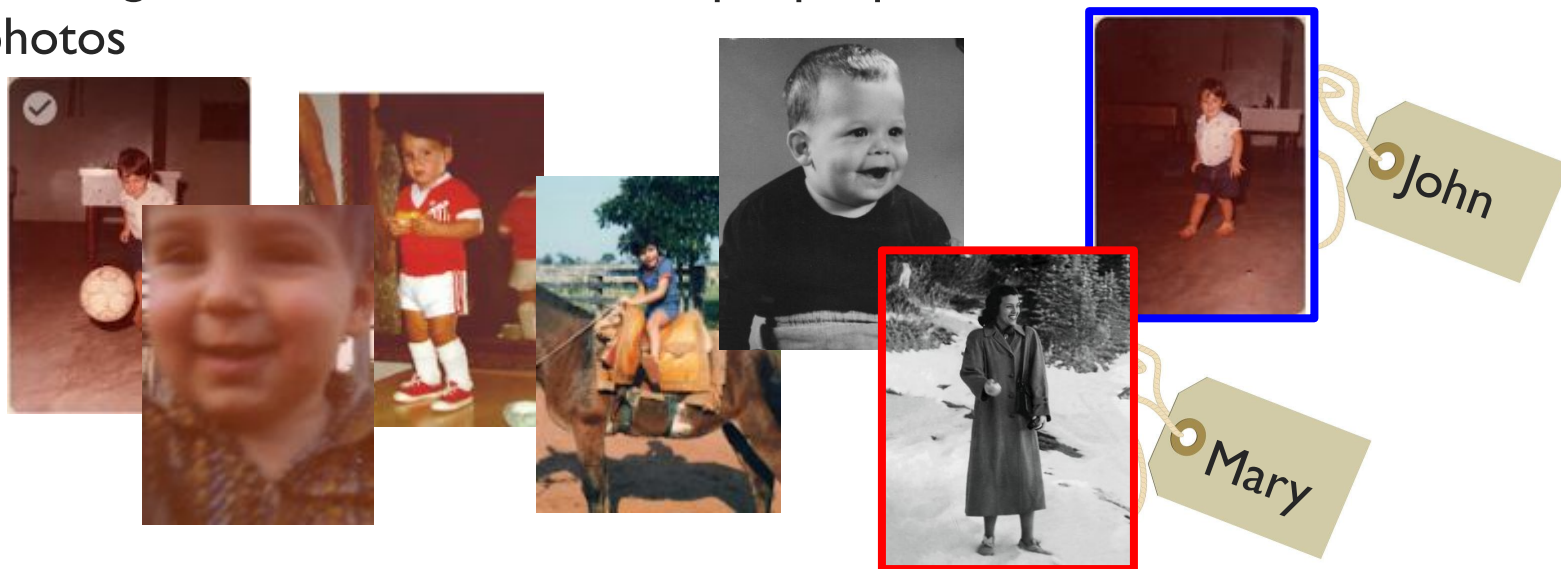
# ML Types: Semi-supervised learning

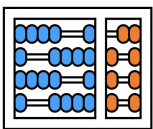


In the **semi-supervised learning**, only a subset of the dataset used to train the model is labeled with the desired/expected solution.

## Example of classification task

- Task: Identify people on photos
- Training data: Photos labeled with people present on it + unlabeled photos





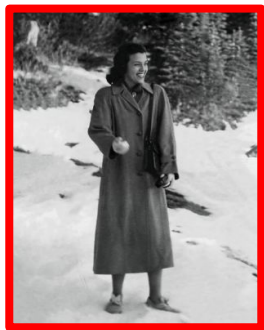
# ML Types: Semi-supervised learning

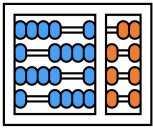


In the **semi-supervised learning**, only a subset of the dataset used to train the model is labeled with the desired/expected solution.

## Example of classification task

- Task: Identify people on photos
- Training data: Photos labeled with people present on it + unlabeled photos



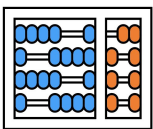


# Machine Learning Overview



## ML Types

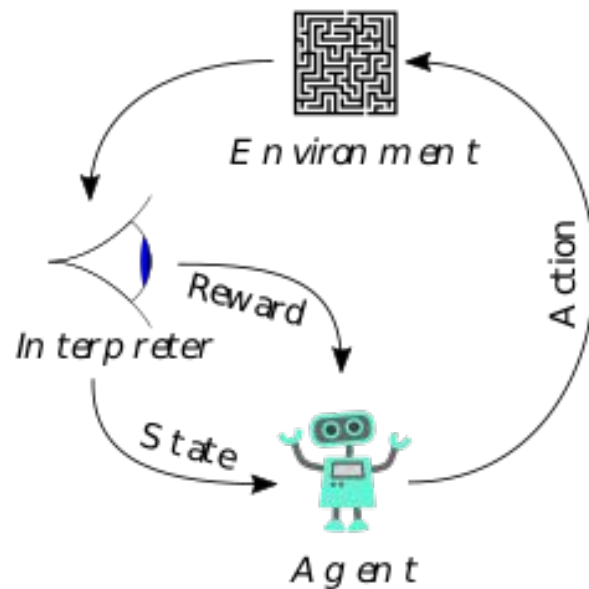
### Reinforcement learning

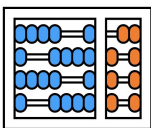


# ML Types: Reinforcement learning



In the **reinforcement learning**, there is no dataset. The learning system (called agent) can observe the environment, select and perform actions, and get *rewards* (or *penalties*). Based on the *reward/penalty* values, the system adjust the policy (learning) to get most of the *reward* over time.



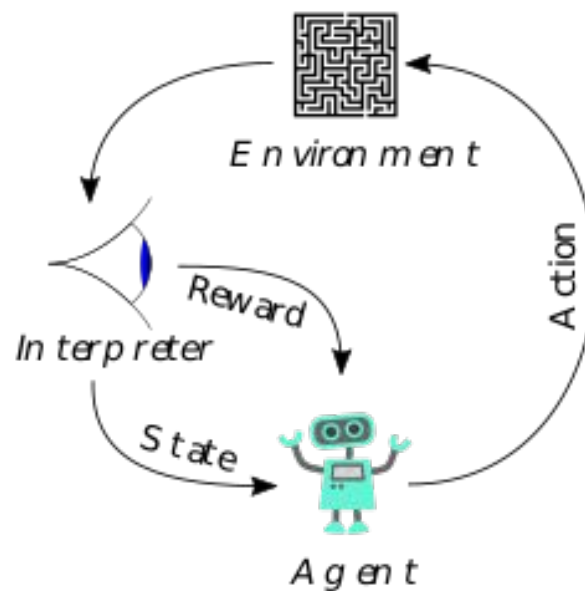


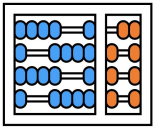
# ML Types: Reinforcement learning



In the **reinforcement learning**, there is no dataset. The learning system (called agent) can observe the environment, select and perform actions, and get *rewards* (or *penalties*). Based on the *reward/penalty* values, the system adjust the policy (learning) to get most of the *reward* over time.

Training of an agent is a process of **trial** and **error**.





# ML Types: Reinforcement learning

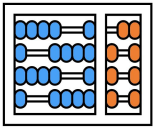


In the **reinforcement learning**, there is no dataset.

Examples:

- Self-driving cars,
- Game playing robots (Go, Chess, Super Mario World, ...)

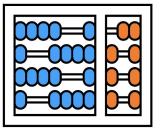
Used often to train  
control/decision  
making systems



# Machine Learning Overview



## ML Types Batch and Online Learning



# ML Types: Batch and Online Learning



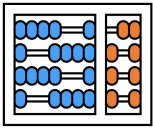
## Batch learning:

- The system is incapable of learning incrementally - it must be trained using all the available data.
  - The system is not capable of adapting to new data on the fly (e.g., in production) - it requires retraining using the whole set.
  - Dataset size may be a challenge

## Online learning:

- The system can be trained incrementally, one sample at a time or by small groups called mini-batches.
  - Can be used to train systems on huge datasets!

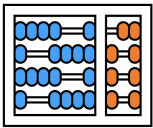




# Machine Learning Overview



## ML Types Generalization Approach



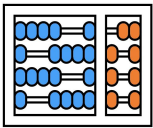
# ML Types: Generalization Approach



ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on new instances (not seen on training).

Two main approaches to generalizations:

- Instance-based learning; and
- Model-based learning



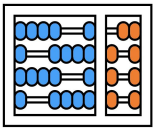
# ML Types: Generalization Approach



ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on new instances (not seen on training).

## Instance-based learning:

- Work by comparing new data points to known data points!
  - Ex: Check whether new emails are similar to emails that had been flagged as SPAM.



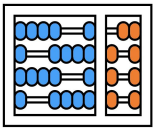
# ML Types: Generalization Approach



ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on new instances (not seen on training).

## Model-based learning:

- Work by detecting patterns in the training data and building a predictive model!



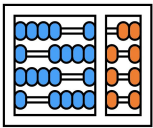
# ML Types: Generalization Approach



ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on new instances (not seen on training).

Example: Orange vs Lemon





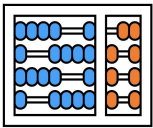
# ML Types: Generalization Approach



ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on new instances (not seen on training).

## Instance-based learning



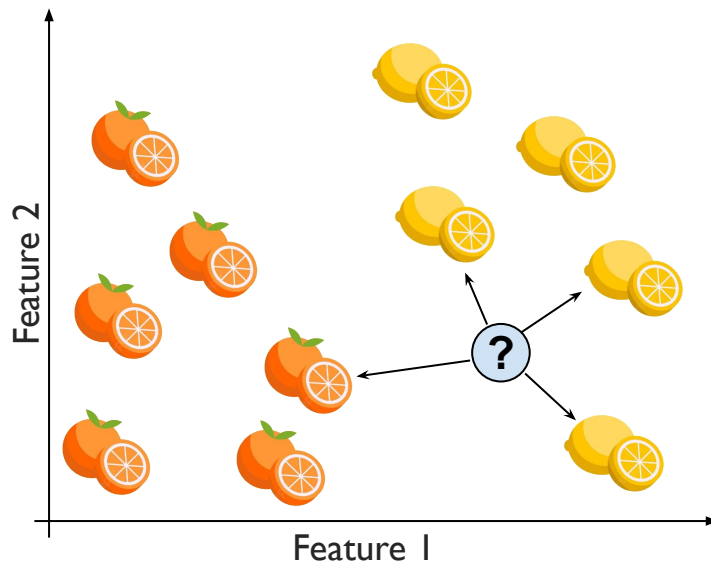


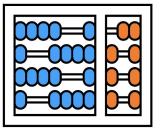
# ML Types: Generalization Approach



ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on new instances (not seen on training).

## Instance-based learning



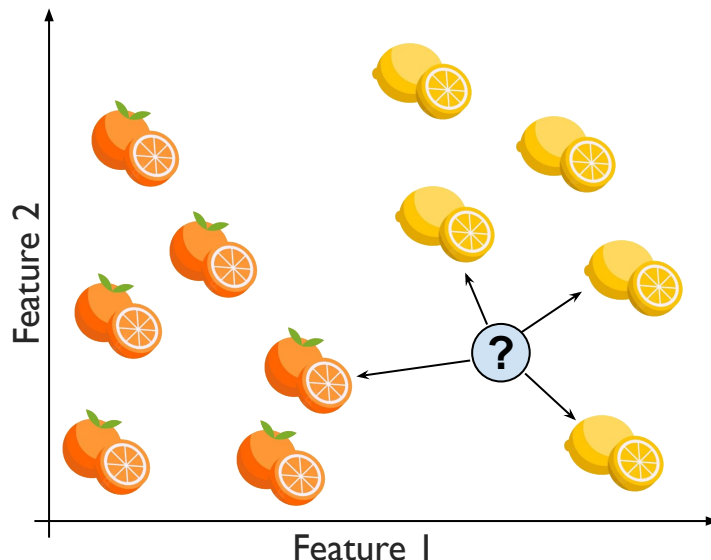


# ML Types: Generalization Approach



ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on new instances (not seen on training).

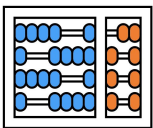
## Instance-based learning



## Model-based learning





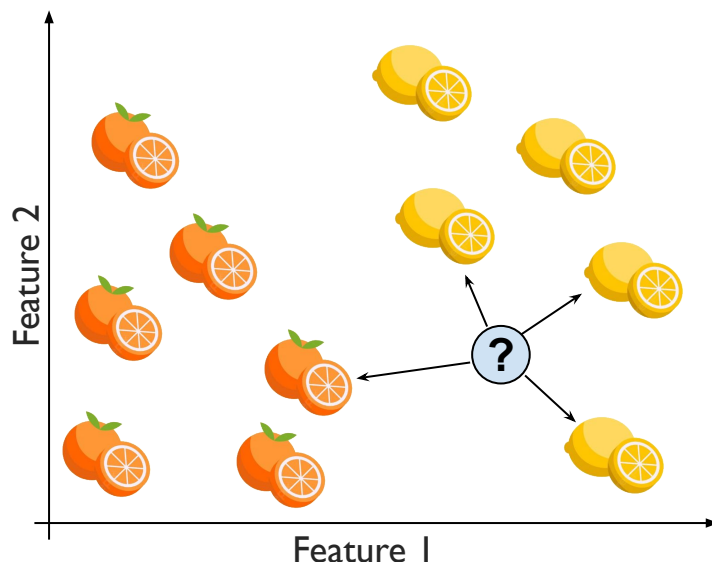


# ML Types: Generalization Approach

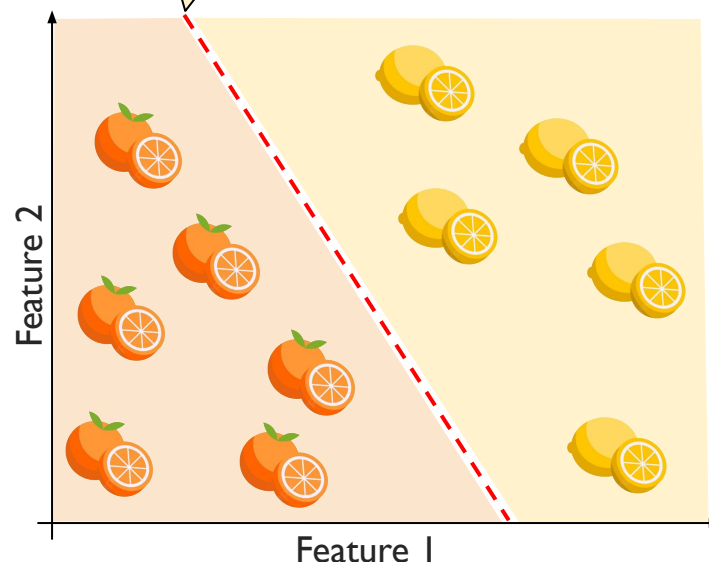


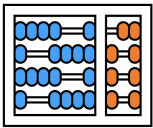
ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on *test* data (not seen on training).

## Instance-based learning



## Model-based learning



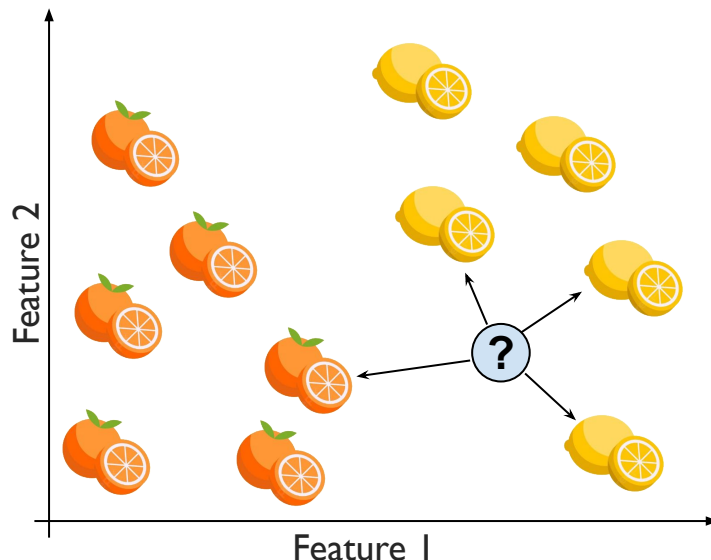


# ML Types: Generalization Approach

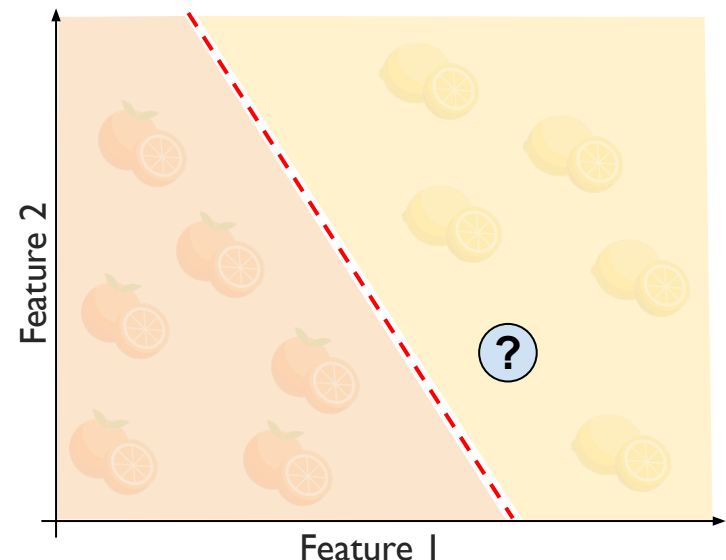


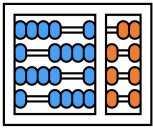
ML models can be classified by how they *generalize*, *i.e.*, how they perform predictions on new instances (not seen on training).

## Instance-based learning



## Model-based learning



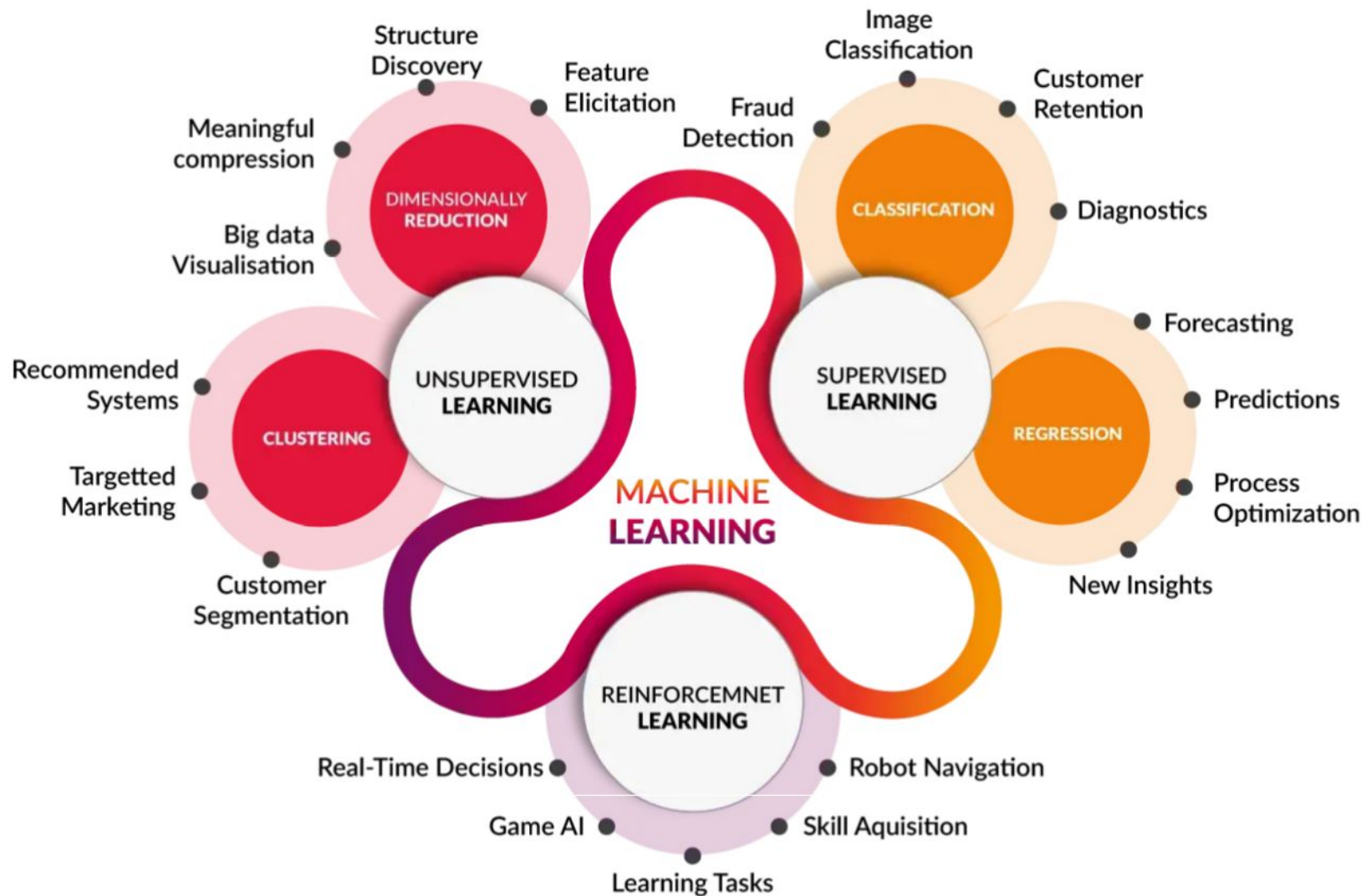


# ML Types: Summary

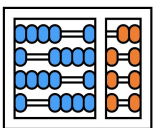


## Summary

# ML Definition and Types: Summary



<https://towardsdatascience.com/coding-deep-learning-for-beginners-types-of-machine-learning-b9e651e1ed9d>



# ML Definition and Types: Summary



## ML Definition by ML :-)

**Edson Borin:**

What is machine learning?

**ChatGPT:**

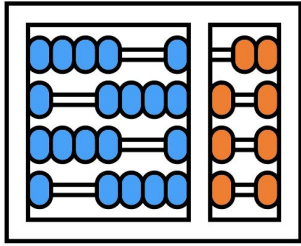
Machine learning is a **branch of artificial intelligence** that focuses on developing **algorithms** and computational models that can **learn and adapt from data**. This **allows systems to automate complex tasks without the need for explicit programming**.

There are several types of machine learning, including supervised learning, unsupervised learning, reinforcement learning, and others.

Supervised learning is when the algorithm is trained with a set of labeled data, meaning, where the output is already known, and it tries to learn to predict the output based on the input data. In unsupervised learning, the algorithm is trained with unlabelled data, and it tries to find patterns and structures in the data. Reinforcement learning is when the algorithm is trained with a series of actions and rewards, and it tries to find the action that maximizes the reward.

Machine learning is widely used in various areas, such as image recognition, natural language processing, speech recognition, prediction, among others.





**Instituto de  
Computação**

UNIVERSIDADE ESTADUAL DE CAMPINAS



# Capacitação profissional em tecnologias de Inteligência Artificial

## **Machine Learning Overview**

**Prof. Edson Borin**

<https://www.ic.unicamp.br/~edson>

Institute of Computing - UNICAMP