Powered by Plain Concepts

#STechDay2021

Glovo

# SINGULARITY **TECH DAY** 2021

The era of AI and Cognitive Services

## Machine learning for on-demand delivery services

Maxim Khalilov, Head of Data Science



## Agenda

01 / The Glovo story

02 / Machine learning at Glovo

03 / Dispatching use case for machine learning

04 / Machine Learning metrics



#STechDay2021

## The Glovo story



## Aiming to connect citizens with possibilities

#STechDay2021

Born in Barcelona by Oscar Pierre in 2015, when he was 22

Let's

GIVE EVERYONE EASY ACCESS TO ANYTHING IN THEIR CITY





#STechDay2021

# Machine learning at Glovo

### Who are the ML users?



#### #STechDay2021



#### **DURING AN ORDER**

Optimal assignment and delivery

Better incentives for difficult orders

Incident resolution & LiveOps Assistance

#### AFTER AN ORDER

Customer/Courier/Partner churn

Customer lifetime value

After-sale service

### Tooling

SINGULARITY TECH DAY 2021

- We have standardised libraries and pipelines contained in our Machine Learning Platform.
- We train and serve our models remotely, on **AWS** instances.
- Training, experimentation, deployment has historically been handled using **Jenkins** pipelines.
- More recently we have started using **Argo** for workflow handling.
- We also started deploying our models using **Sagemaker**, but we migrated to using **Kubernetes and Spinnaker**.
- We monitor the health of our systems (latencies/timeouts/errors) using **Datadog**.
- For Data Quality and monitoring, we use **Great Expectations** and **Mona**.



#### #STechDay2021

#STechDay2021

## Dispatching use case for Machine learning



## Who is (Mr.) Jarvis?



## Jarvis is making hundreds of complex decisions every second

- Who's the best courier for each order?
  - Impacts delivery time & CPO (Cost per order)
- When do we notify a partner about a new order?
  - $\circ$  Too early and food will get cold.
  - Too late and the courier will have to wait!
- What store address to select?
  - If there are multiple stores for a given partner, which one fulfills the order?
- Partner and Customer ETAs



## **Dispatching process**



#### #STechDay2021





1 .....



## **Estimated time of arrival**

![](_page_16_Figure_3.jpeg)

![](_page_16_Picture_4.jpeg)

#### ETA problem:

- Challenging: minimal information available, real-time
- **High Impact:** Key factor in customer decision when placing an order

#### ETA solution: Machine Learning + Business insights

![](_page_16_Figure_9.jpeg)

![](_page_17_Figure_2.jpeg)

-

![](_page_18_Picture_1.jpeg)

#### **Key Objective**

Synchronize **courier arrival to the restaurant** and **restaurant finishing preparing the order**.

![](_page_18_Picture_4.jpeg)

Estimated Time of Arrival to Pickup

Challenges: Several components such as travel times, waiting times, etc.

![](_page_18_Picture_7.jpeg)

Estimated Time to Prepare

#STechDay2021

Challenges: Only partial information available

![](_page_19_Figure_0.jpeg)

#STechDay2021

ETP

![](_page_20_Figure_3.jpeg)

## **Online post-processing**

Receives the waiting time directly in production and learns dynamically in production

![](_page_21_Figure_4.jpeg)

![](_page_22_Figure_2.jpeg)

#STechDay2021

### **Courier models**

![](_page_23_Figure_3.jpeg)

## **Courier models: multiple features**

![](_page_24_Picture_3.jpeg)

- Flight distance
- Coordinates
- Manhattan distance
- Bearing
- etc.

![](_page_24_Figure_9.jpeg)

- Time of the day
- Day of the week
- Hour
- Month
- etc.

![](_page_24_Figure_15.jpeg)

![](_page_24_Picture_16.jpeg)

- Orders per shift
- Time difference between acceptance and start/end of the shift
- # of orders delivered over time
- Flight speed over last X orders

#STechDay2021

# Machine learning metrics

![](_page_26_Picture_0.jpeg)

## **ML in prod: metrics**

![](_page_26_Picture_3.jpeg)

![](_page_26_Picture_4.jpeg)

![](_page_26_Picture_5.jpeg)

![](_page_26_Picture_6.jpeg)

General

#### Performance

- Latency
- Availability
- Error: MAE, RMSE
- Accuracy, Precision/Recall
- Primary (updated at each iteration)
- Sanity (fixed for a long period of time)
- Robustness (generic)

#### Data shift

- Feature statistics and its distance from the training data
- Histogram intersectionPopulation
- Stability Index
- Novelty detection

#### Controlled online experiments

- A/B testing
- Switchbacks
- MABs
- Other

#STechDay2021

![](_page_27_Picture_2.jpeg)

## We are hiring!

singularity **TECH DAY** 2021

## Thank you!

plain concepts

![](_page_28_Picture_2.jpeg)

![](_page_28_Picture_3.jpeg)