



Topic modeling and importance classification of warning messages (NOTAMs) for SWISS international airlines

Jean Coupon
October 17, 2018

Data Science / Batch 4 / July 30 to October 19, 2018

Who am I?

- Ph.D. in Astrophysics (2009)
- 4 years of postdocs in Asia (Japan and Taiwan)
- research associate at the University of Geneva since 2013
- Seeking new challenges in industry to work on business-driven (“real-life”) projects

Warning messages for pilots (NOTAMs)

- NOTAM = Notice for Airmen
- Short text with 2 parts: **codes** + **plain English**
- Created by Aeronautical authorities, gathered globally and provided to airlines
- Describe any abnormal condition affecting flights (but weather): obstacle, instrument malfunctions, prohibited airspace (military exercises)

B0821/18 NOTAMN

Q) LSAS/QOLAS/V/M/E/000/035/4652N00837E005

A) LSAS B) 1807161442 C) 1807232359 EST

E) OBST LGT 2.5KM W ALTDORF (ANTENNA ATTINGHAUSEN) 465157N0083644E

Challenge

- NOTAM workflow at SWISS:
 1. receive messages daily,
 2. suppress no-fly zones (remove only 1%),
 3. classify according to importance (NOTAM officer)
 4. incorporate to flight briefing (flight dispatchers)
 5. hand over to pilots
- 3000 messages received per day
- All done manually

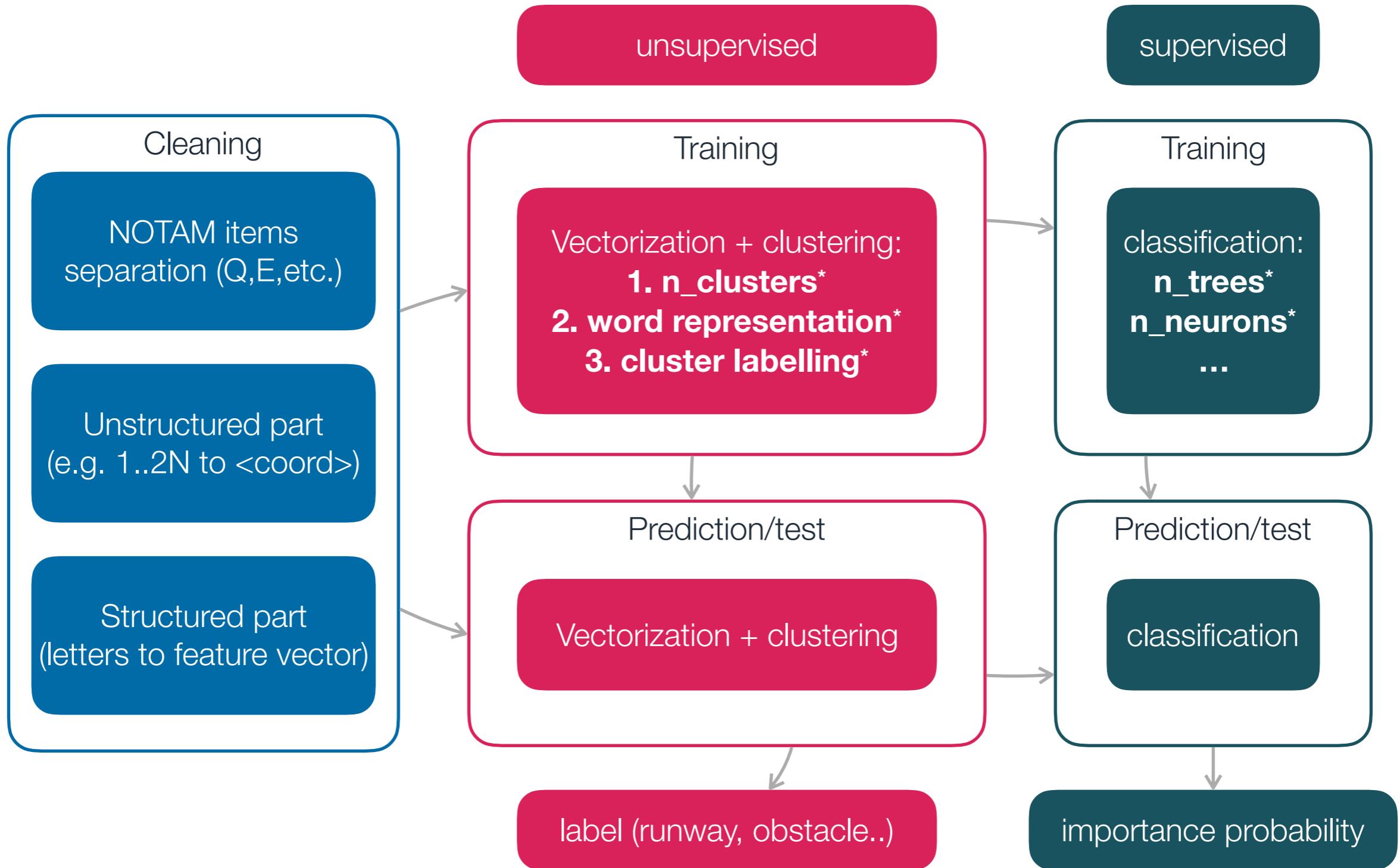
Goals

- Can we assist the workflow with automated processes?
- Two goals:
 - labelling the messages using language patterns in the data
(unsupervised machine learning)
 - assigning a probability of importance
(supervised machine learning)
- Millions of classified NOTAMs over decades
- project: develop a software pipeline to label and assign probability for each NOTAM. Used a sample of 100,000 human-classified NOTAMs

Goals

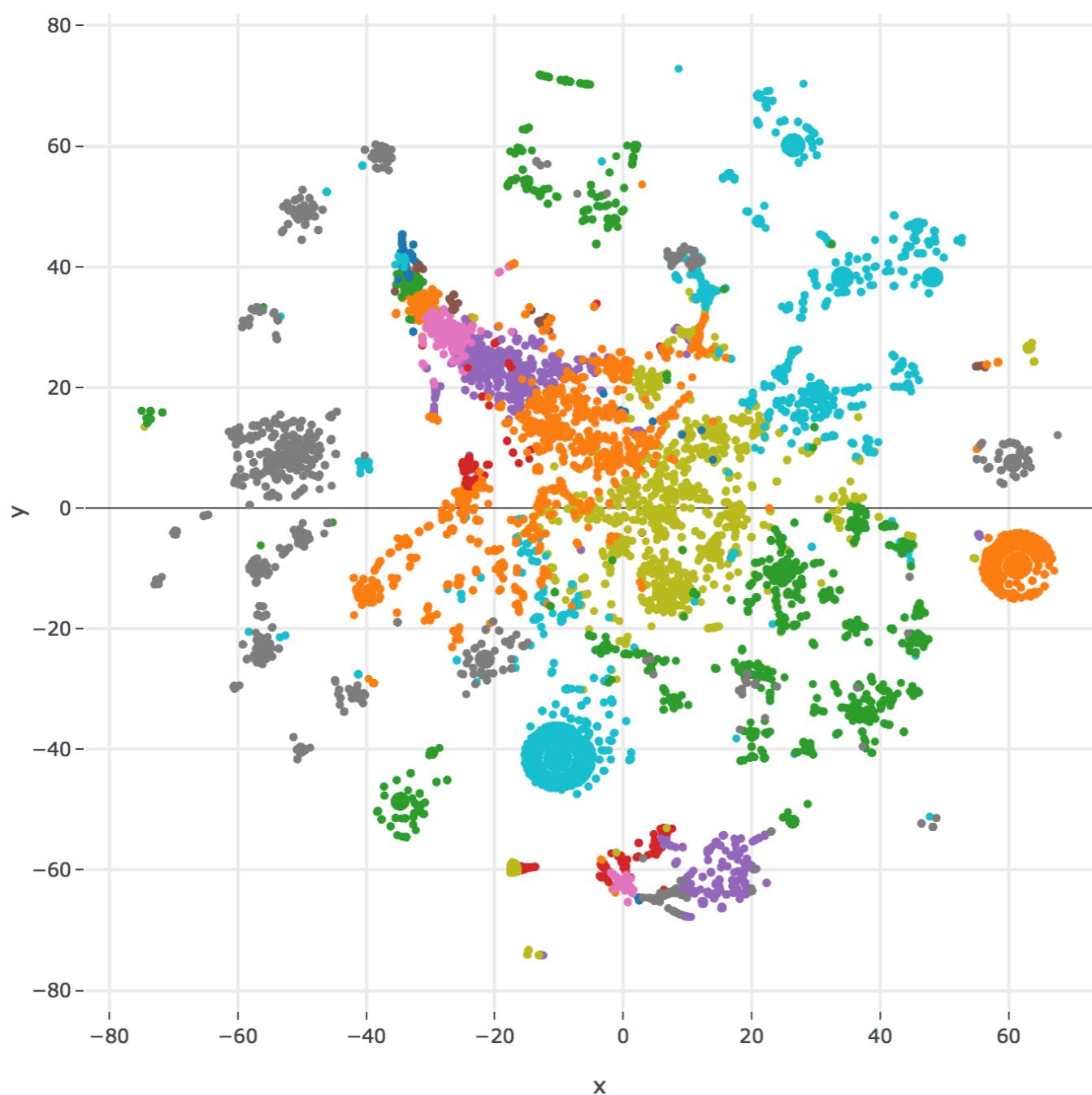
Objective: help everyone in the process to make **informed decisions**, save time with obvious NOTAMs (save money), spend more time on uncertain cases (safety increase)

NOTAMs modeling workflow



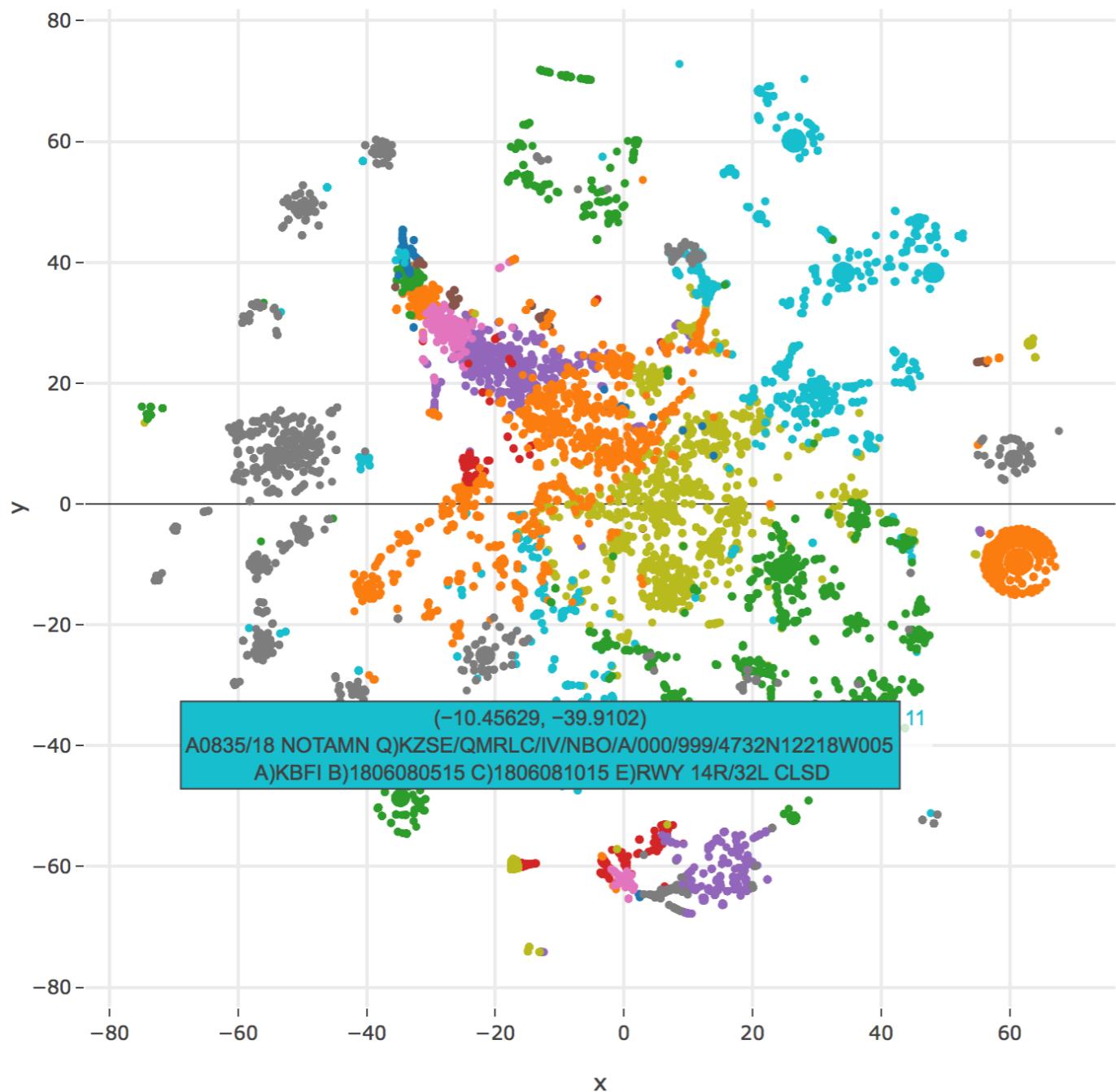
(*model tuning)

Labelling the NOTAMs with unsupervised learning



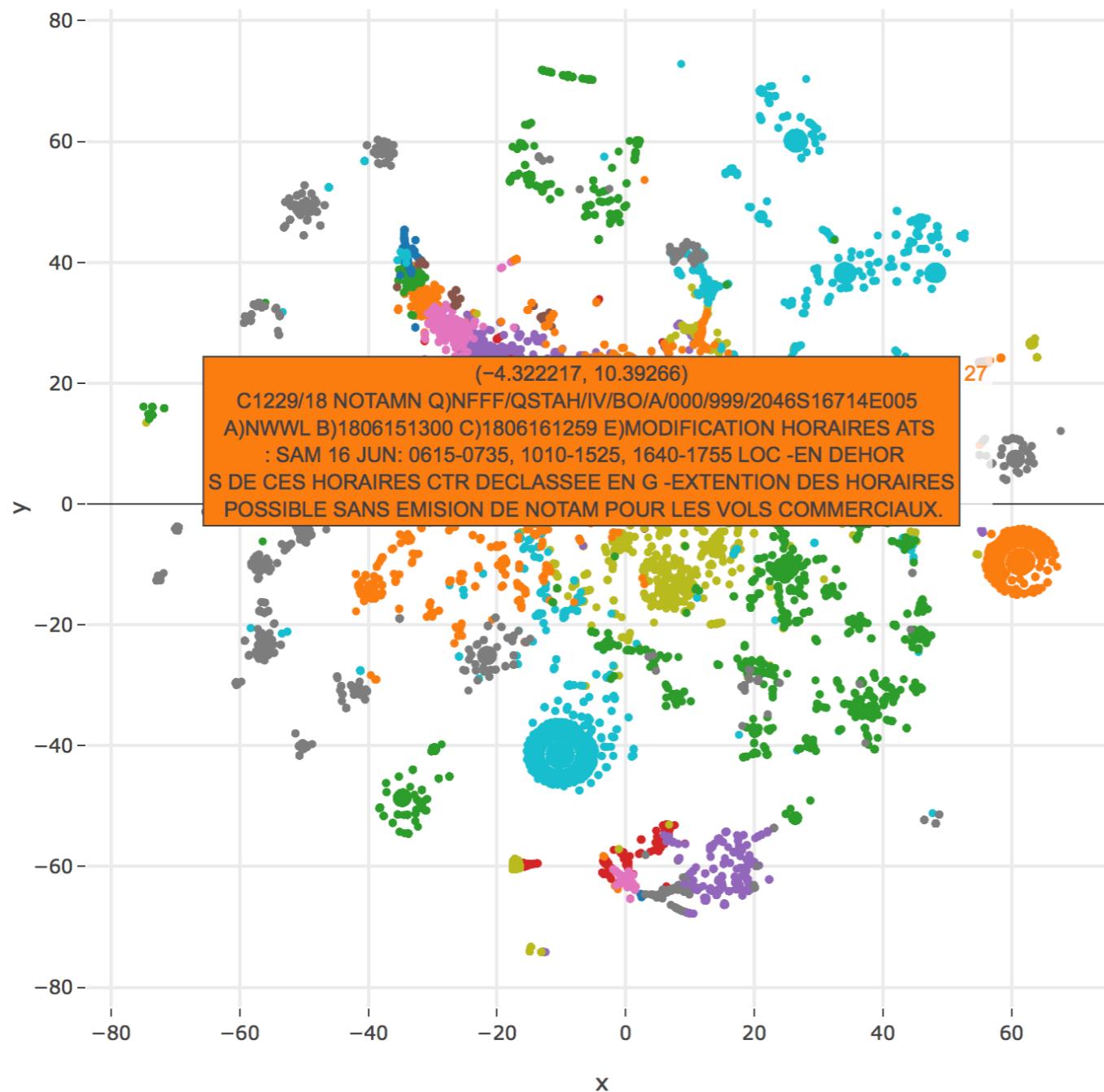
- Similar NOTAMs are grouped together
- Highly tuneable unsupervised ML model
- User can manually decide the number of “clusters”
- And then do the labelling

Important NOTAMs



Closed runway
(very important)

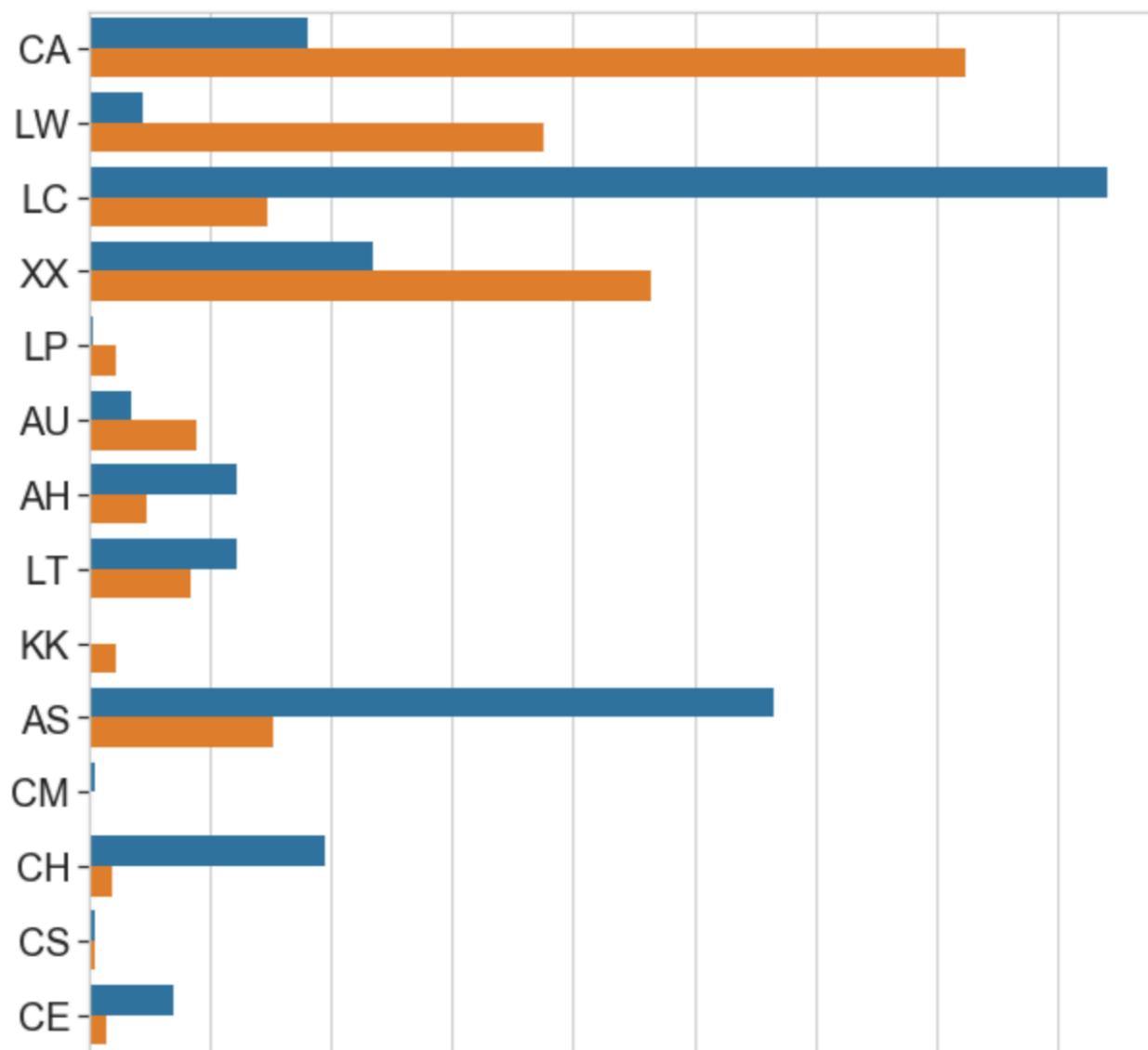
Uncertain NOTAMs



Message in French
(important?)

Importance classification

- Use codes



important

suppressed

Importance classification

```
Precision: 0.8907, recall: 0.9064, accuracy: 0.9092
      precision    recall   f1-score   support
          0         0.92      0.91      0.92      8368
          1         0.89      0.91      0.90      6664
avg / total         0.91      0.91      0.91     15032
```

Future

- Use more NOTAM to refine the model
- Tune the model to improve labelling
- Get user feedback (pilots, NOTAM officers)

Thank you!

Contact info