Marco Castelluccio

Tutor: Carlo Sansone – co-Tutor: Annalisa Verdoliva

XXXI Cycle - I year presentation

## Improving software quality using machine learning

Background

- MSc with honour in Computer Engineering at University of Naples Federico II
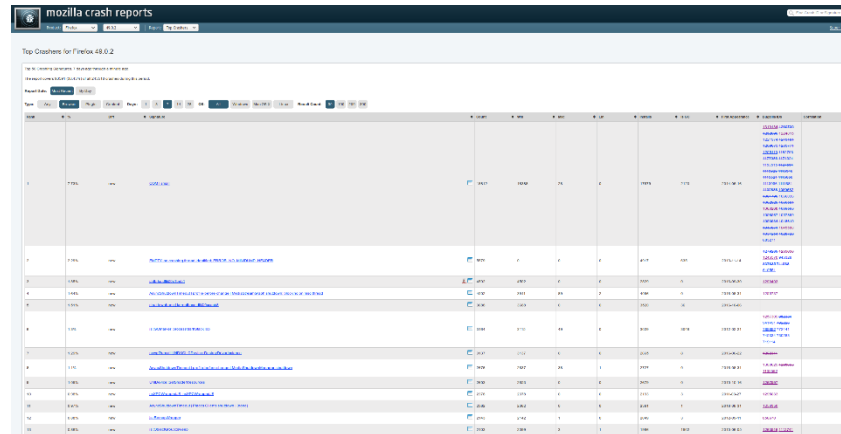
- Software Engineer at Mozilla

Automating the understanding of groups of crash reports

- Many studies focused on improving the bucketing of crash reports. My focus is on how to automatically describe the buckets' properties in the most interesting way for developers.
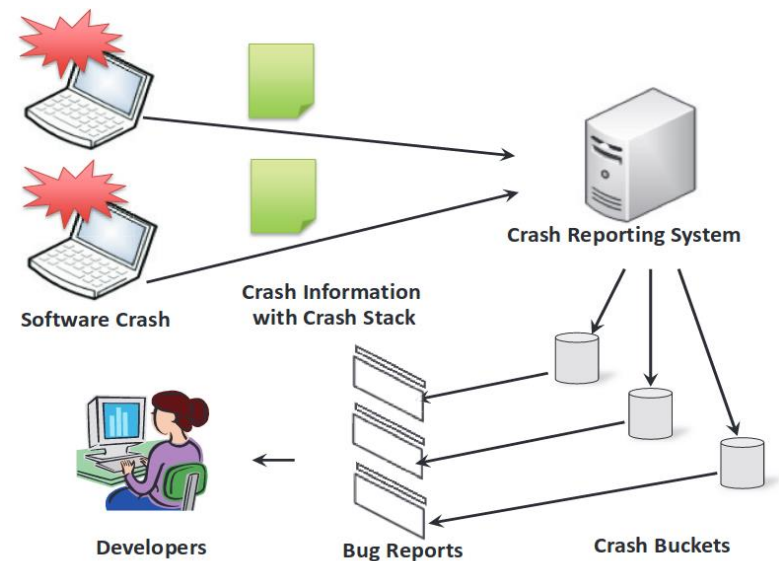
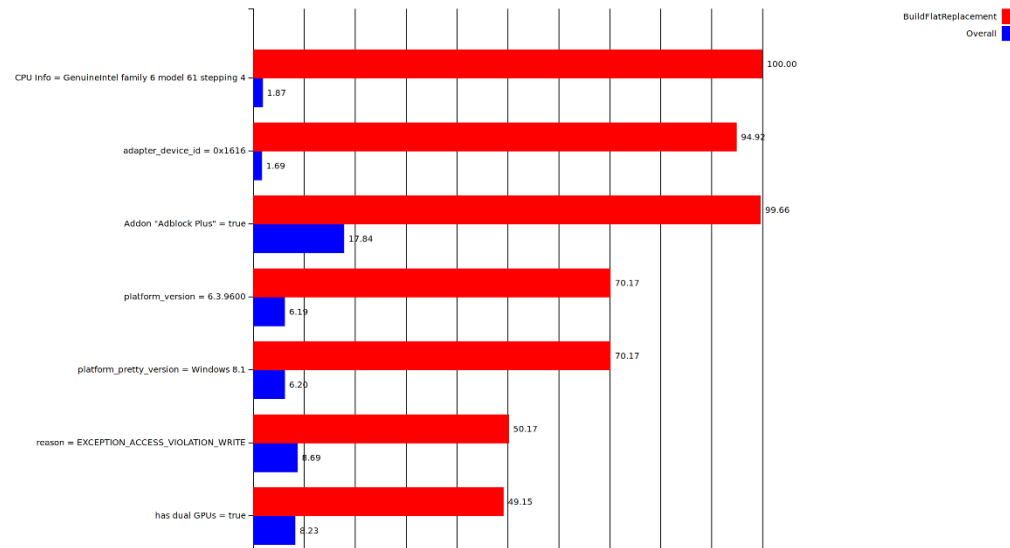# Automating the understanding of groups of crash reports

- For the Mozilla Firefox product, there are more than 1 million crash reports per week.

- Every crash report contains, other than the crash stack, hundreds of attributes (e.g. CPU architecture, graphic card driver version, list of installed addons, etc.).


Marco Castelluccio

Automating the understanding of groups of crash reports

- Understanding what makes a crash group meaningfully different than other groups is very often useful for debugging (sometimes even enough for fixing the crash, e.g. by blocklisting a certain gfx card).

Marco Castelluccio

5

# Automating the understanding of groups of crash reports

- The algorithm is based on the STUCCO data mining algorithm.
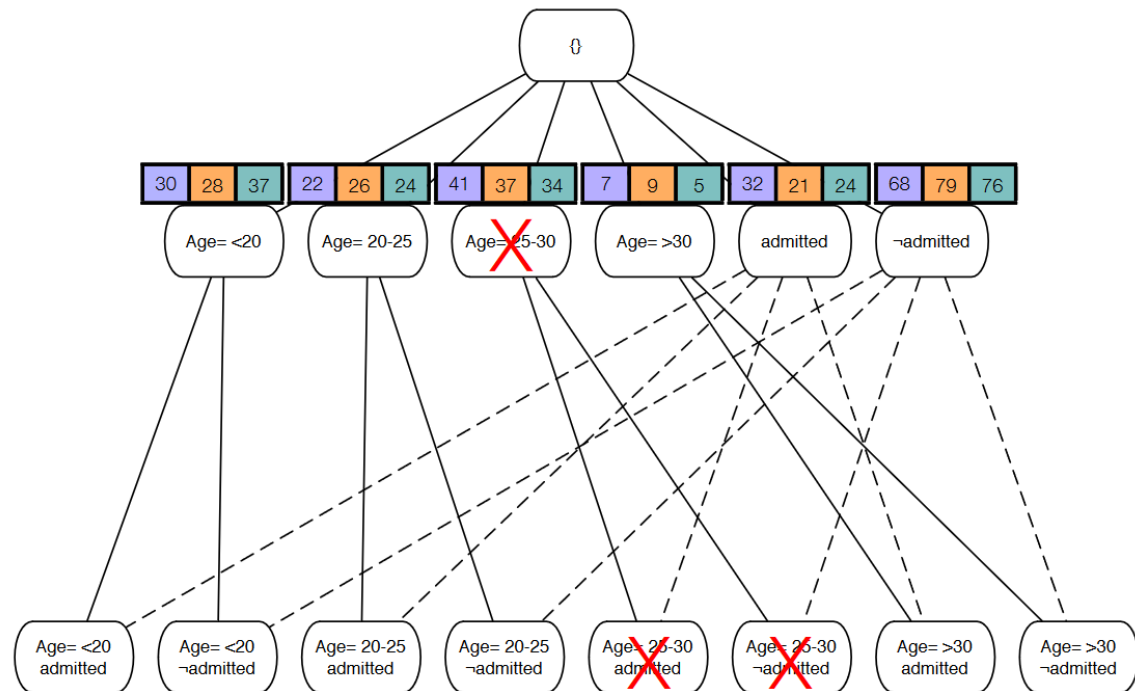
Empirical study of uplifts (backports) in Mozilla Firefox

- Collaboration with the École Polytechnique de Montréal.

- Uplifts are critical changes for a software, primarily because they have much less time to stabilize.

- The aim is to understand their properties vs normal changes; understand which uplifts introduced bugs and why; with the ultimate goal of building a model to predict the riskiness of an uplift.

Marco Castelluccio

Empirical study of uplifts (backports) in Mozilla Firefox
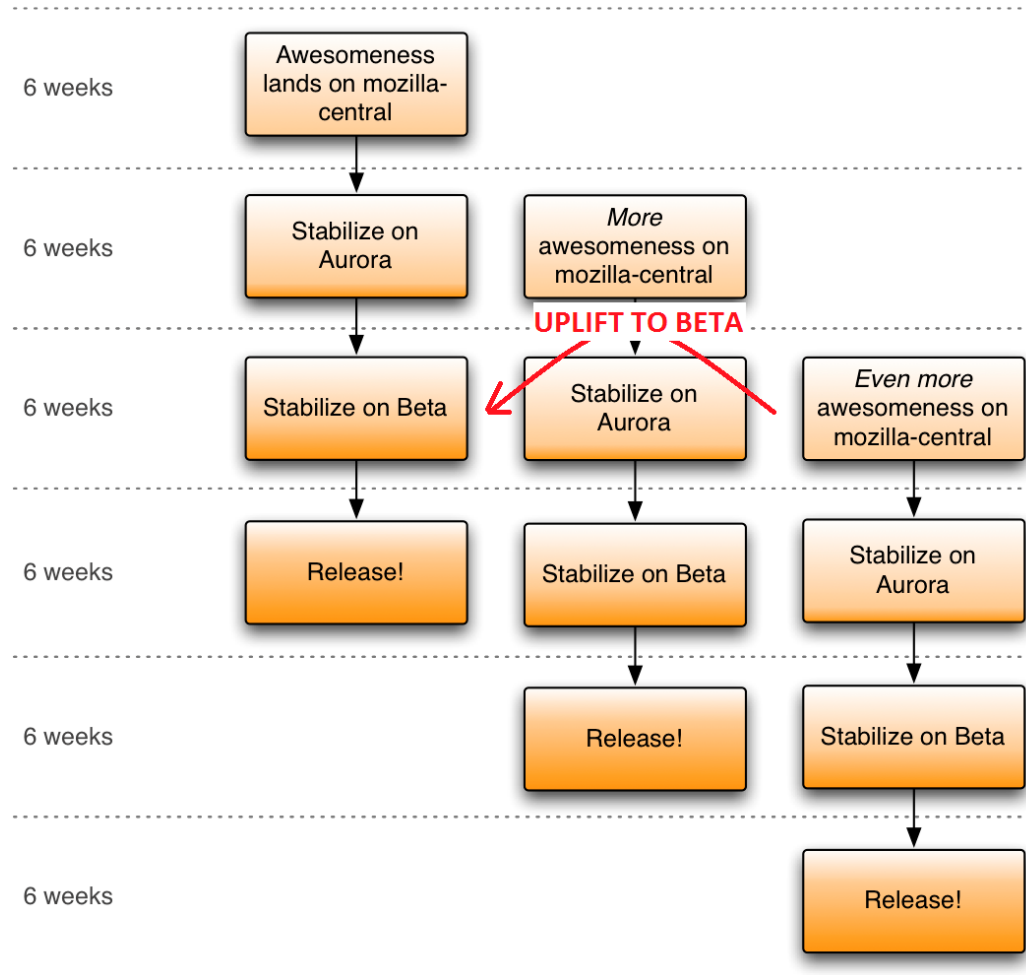
Publications

- M. Castelluccio, G. Poggi, C. Sansone, L. Verdoliva – Land Use Classification in Remote Sensing Images by Convolutional Neural Networks – https://arxiv.org/abs/1508.00092

- M. Castelluccio, G. Poggi, C. Sansone, L. Verdoliva – Training Convolutional Neural Networks for Semantic Classification of Remote Sensing Imagery – JURSE2017 (submitted)

Next Years

Year	Modules	Seminars	Research	Tot.
1	21 (20)	7 (5)	35 (35)	63 (60)
2	(9)	(6)	(42)	57 (60)
3	(0)	(5)	(55)	60 (60)
Tot.	30 (30-70)	18 (10-30)	132 (80-140)	180 (180)

Marco Castelluccio