REDUCING FALSE POSITIVES IN INTRUSION DETECTION
RAPID DETECTION & RESPONSE SERVICE

• F-Secure solution for intrusion detection in corporate IT environments
• Sensors (=program on a client computer) sends events to our backend
• In the real-time processing component of the solution, events are analyzed by machine learning models and rules, triggering detections
• Detections are evaluated by the experts at Rapid Detection Center (RDC)
• RDC decides whether to contact the client or dismiss the detection
• Rules are prone to false positives. Here we aim to predict which rule-based detections could be false positives to prioritize the most likely incidents
CHALLENGES WITH DATA #1

• Underlying distribution changes constantly
  • new or updated rules
  • new or changed clients
  • new or changed applications on clients
  • ...

• Significant class imbalance
CHALLENGES WITH DATA #2

• Significant rule imbalance
  • a lot of rules trigger detections rarely, or not at all

• Different rules have different sets of features

• Assuming IID is questionable
  • if there is a detection from the same host/organization, it is likely that similar detections will occur again soon
CHALLENGES WITH DATA #3

• Detections come in a stream but...
  • ...they are not labelled in the same order as labels come after human processing
  • detections can also be grouped in batches, and then batches are labelled

• Evaluating performance of a classifier becomes non-trivial
  • data leakage is a real issue, with significant impact

• Software development challenges due to data sensitivity/volume.
INITIAL RESULTS:
RESULTS ALREADY LOOK QUITE PROMISING