

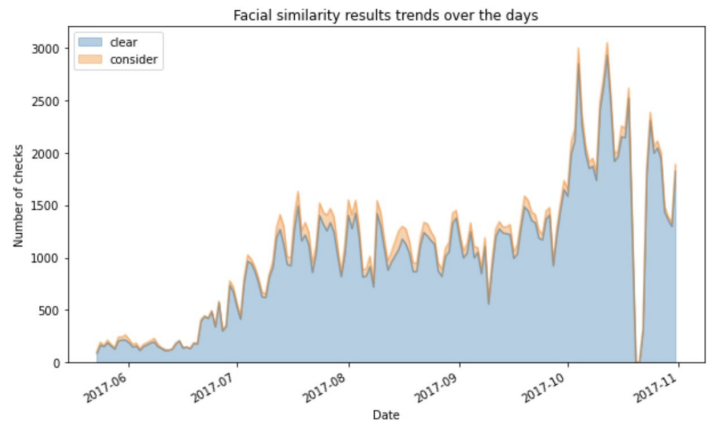
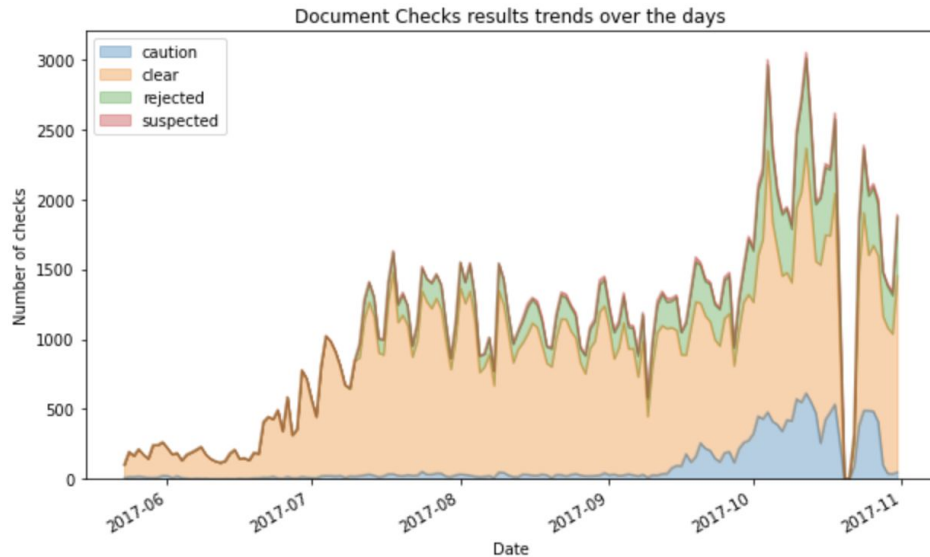
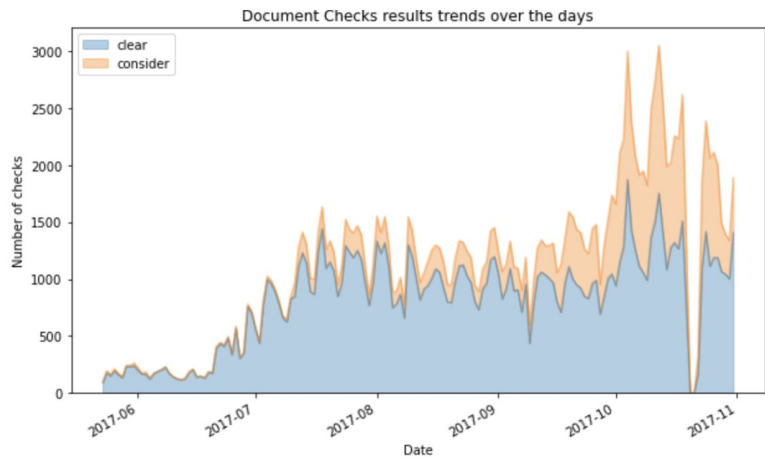


Revolut Analytical Challenge

Timothée Jauffret - 2021-11-28

This presentation can be completed with the following python notebook.

The analysis of results distribution shows Document check process is responsible for the decrease of pass rate. From July 10th "Rejected" in the main cause. From Sept 10 we start having a huge increase of "Caution" sub results



- Document check is responsible for the decrease in pass percentage
- The most common sub-Results are Rejected and Caution
- Caution issue seems to come back to its normal level by the end of October.

In the next slide we will deepdive into each features to detect if there is one responsible for low pass rate

Main reason for failed in rejected cases is due to image quality (94% of the cases)
Main reason for failed in caution cases is due to incomplete image

REJECTED CASES

| | Events | Event pass rate | Number of clear results |
|-----------------------------|---------|-----------------|-------------------------|
| Feature | | | |
| image_integrity_result | 26090.0 | 0.000000 | 0.0 |
| data_validation_result | 3.0 | 0.666667 | 2.0 |
| compromised_document_result | 3.0 | 1.000000 | 3.0 |
| visual_authenticity_result | 0.0 | NaN | NaN |
| data_consistency_result | 0.0 | NaN | NaN |
| data_comparison_result | 0.0 | NaN | NaN |
| police_record_result | 0.0 | NaN | NaN |

Feature pass rate

| | Events | Event pass rate | Number of clear results |
|------------------------------------|---------|-----------------|-------------------------|
| Feature | | | |
| image_quality_result | 26090.0 | 0.064661 | 1687.0 |
| supported_document_result | 25589.0 | 0.934073 | 23902.0 |
| conclusive_document_quality_result | 0.0 | NaN | NaN |
| colour_picture_result | 0.0 | NaN | NaN |

Sub Feature pass rate

94% of rejection is due to image quality

CAUTION CASES

| | Events | Event pass rate | Number of clear results |
|-----------------------------|---------|-----------------|-------------------------|
| Feature | | | |
| image_integrity_result | 15992.0 | 0.159705 | 2554.0 |
| data_comparison_result | 196.0 | 0.173469 | 34.0 |
| visual_authenticity_result | 15970.0 | 0.909956 | 14532.0 |
| data_validation_result | 15584.0 | 0.922164 | 14371.0 |
| data_consistency_result | 9858.0 | 1.000000 | 9858.0 |
| police_record_result | 15589.0 | 1.000000 | 15589.0 |
| compromised_document_result | 10462.0 | 1.000000 | 10462.0 |

Feature pass rate

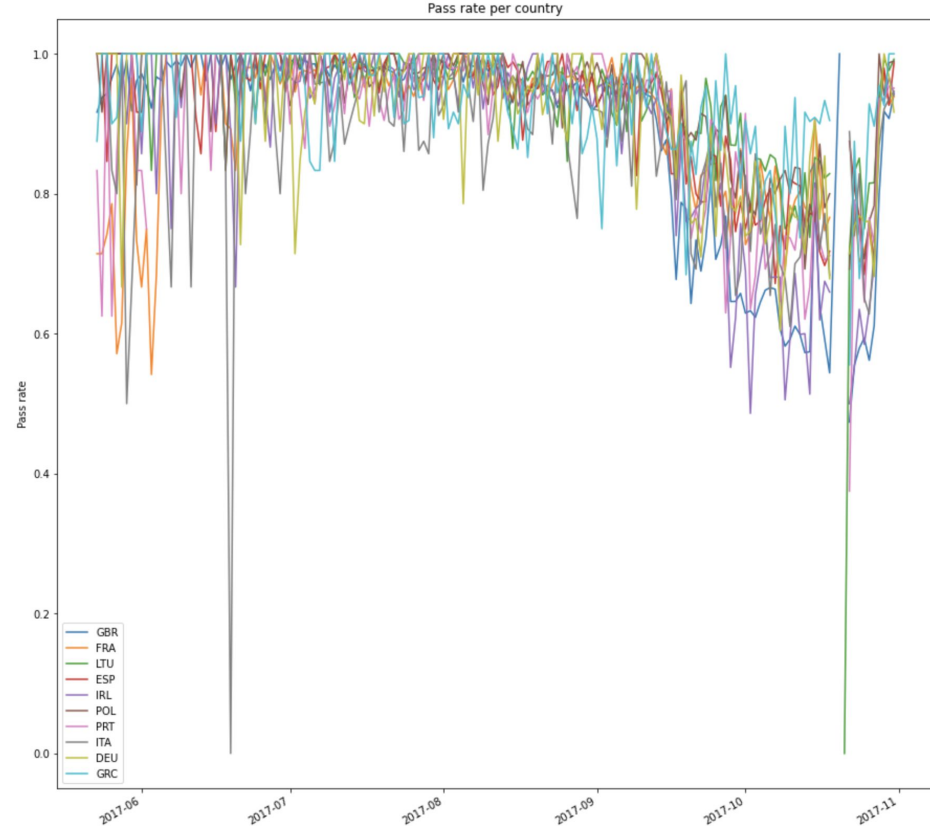
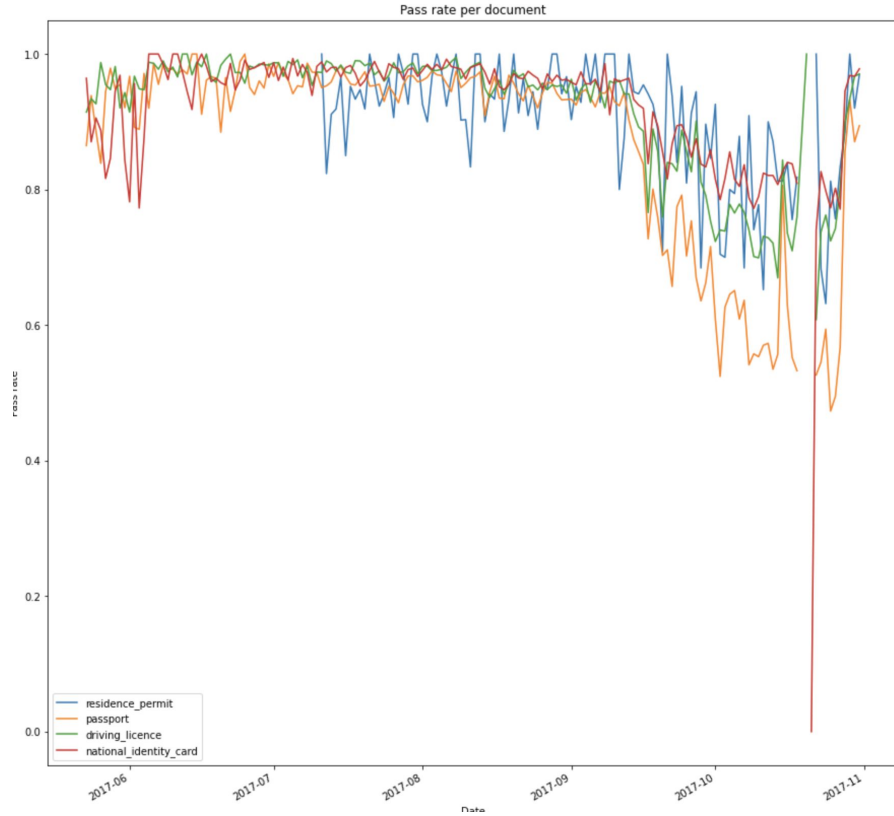
| | Events | Event pass rate | Number of clear results |
|------------------------------------|---------|-----------------|-------------------------|
| Feature | | | |
| conclusive_document_quality_result | 14509.0 | 0.07740 | 1123.0 |
| colour_picture_result | 14509.0 | 0.99552 | 14444.0 |
| image_quality_result | 15992.0 | 1.00000 | 15992.0 |
| supported_document_result | 15992.0 | 1.00000 | 15992.0 |

Sub Feature pass rate (image integrity only)

Main rejection is due to conclusive document quality failed followed by face detection and data validation. We are missing details for data validation

Both issues are related to images. For qualities issues it could come from a specific document or countries. We have to analyse user if there is any relation between user properties and pass rate.

The analysis of trends pass rate per document and per country shows that the decrease is generalized across countries and document type.



Conclusion



- Users properties don't impact in trends. The only relation that we see is that as the user base is growing pass percentage is decreasing.
- Rejected cases due to low quality images increase a lot with the growth of the company. It could be explained by more non used-to users, users with low quality phones etc.
 - Add a proper quality check before sending to Veritas in order to give user direct feedback and avoid double processes
 - Give more details on how to take a good picture of the document
- Caution cases increase a lot during highest peak of new customers and then seems to go back to normal rate.
 - Explore any API connection from our side our services could be impact by highest peak of traffic
 - Same thing for Veritas. Organize a meeting with their engineers to detect if their API can support such flow
 - Consider other verification organization in case Veritas can't improve their capacity, to have a automatic second check for Caution issues