

# Customer Success

Εφαρμογή για την οπτικοποίηση και την ανάλυση των παραπόνων των επιβατών του ΟΑΣΑ (Σύστημα εξυπηρέτησης πολιτών και επιβατών νέας γενιάς)



# Περιγραμμά Παρουσίασης

- Σκοπός και Περιγραφή του Έργου
- Σύνθεση Ομάδων
- Δεδομένα που χρησιμοποιήθηκαν και Προεπεξεργασία Δεδομένων
- Προτάσεις βελτίωσης διαδικασιών καταχώρησης και προτυποποίησης παραπόνων
- Μηχανισμοί που υλοποιήθηκαν
- Επόμενα Βήματα

# CustomerSuccess

- Ο σκοπός του έργου είναι η ανάπτυξη ενός συστήματος εξυπηρέτησης πολιτών και επιβατών για την αντιμετώπιση προβλημάτων. Οι βασικοί στόχοι του έργου είναι οι εξής:
  - Η συλλογή και αποθήκευση με αυτοματοποιημένο, μεθοδικό και προτυποποιημένο τρόπο των δεδομένων που παράγονται από τους πολίτες ή επιβάτες.
  - Η ανάλυση των δεδομένων, με χρήση της επιστήμης ανάλυσης δεδομένων και τεχνικές μηχανικής μάθησης και τεχνητής νοημοσύνης
  - Η παραγωγή πληροφοριών επιχειρησιακής ευφυίας (business intelligence) προς την διοίκηση, τους συμβούλους και τους υπεύθυνους λήψης αποφάσεων σε όλα τα επίπεδα
  - Η βελτιστοποίηση της πληροφορίας και των καναλιών επικοινωνίας προς τους πολίτες και τους επιβάτες

# CustomerSuccess team - ΟΑΣΑ

- Νίκος Γκοργκόλης - Προϊστάμενος Πληροφοριακών Συστημάτων & ΤΠΕ
- Χριστίνα Βάρλα - Διευθύντρια Μάρκετινγκ και Επικοινωνίας
- Ανάλυση επιχειρησιακών δεδομένων για τη διαμόρφωση της πρότασης
- Διάθεση dataset 'Δεδομένα Παραπόνων Επιβατών' ανά μήνα για το έτος 2022
- Διάθεση συγκεντρωτικών αποτελεσμάτων και γραφημάτων

# CustomerSuccess team -Αρχιμήδης

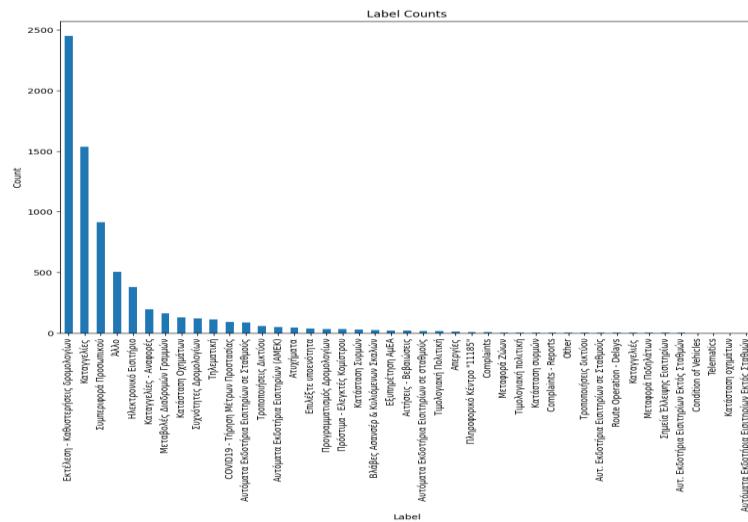
- Γεώργιος Βιτζηλαίος
  - Ιωάννης Δαλιάνης
  - Ματίνα Δέμτση
  - Κωνσταντίνος Χαλιώτης
  - Δημήτρης Πανταζάτος
- Σχεδιασμός και Υλοποίηση Περιβάλλοντος Οπτικοποίησης
  - Υλοποίηση Μηχανισμού Επειγ. Παραπόνων και επεξηγησιμότητας των μοντέλων
  - Υλοποίηση Μηχανισμού ανάλυσης συναισθημάτων
  - Υλοποίηση μηχανισμού κατηγοριοποίησης παραπόνων και επεξεργασία και καθαρισμός των δεδομένων

# Δεδομένα που χρησιμοποιήθηκαν

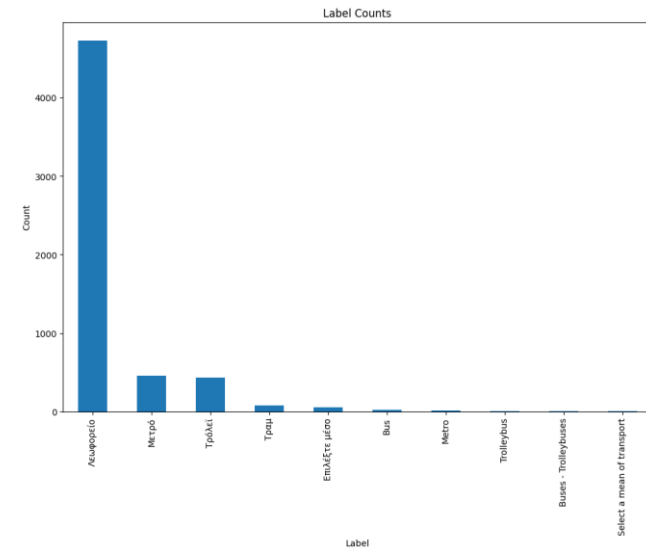
- ▶ Τα δεδομένα που μας δόθηκαν από τον ΟΑΣΑ είναι τα εξής:
  - ▶ Παράπονα ανα μήνα και κατηγοριοποίηση τους για το 2022 (raw και επεξεργασμένα)
  - ▶ Συγκεντρωτικά αποτελέσματα και γραφήματα που έχουν στην παρούσα φάση
- ▶ Πλήθος παραπόνων που χρησιμοποιήθηκαν για την ανάλυση: 7138

# Προεπεξεργασία Δεδομένων

Κατανομή Παραπόνων με βάση τις Αρχικές Κατηγορίες

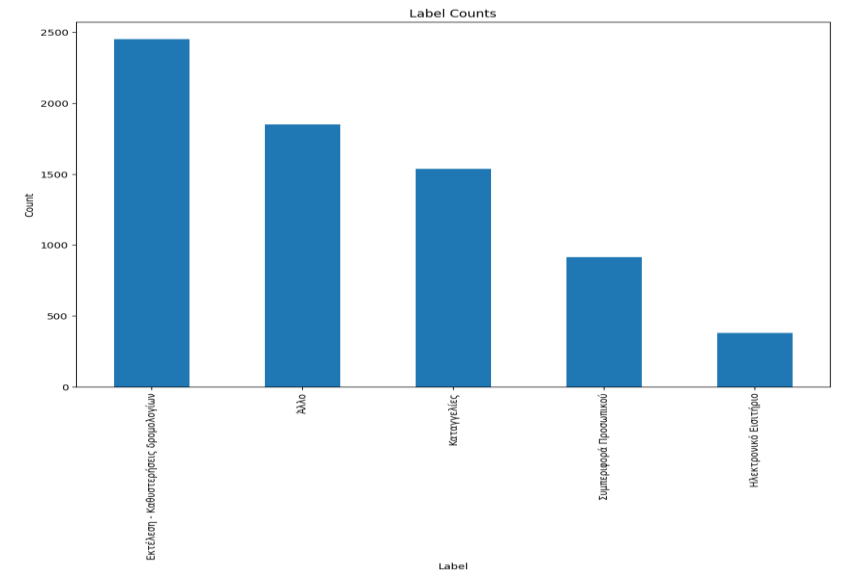


Κατανομή Παραπόνων με βάση το Μέσο



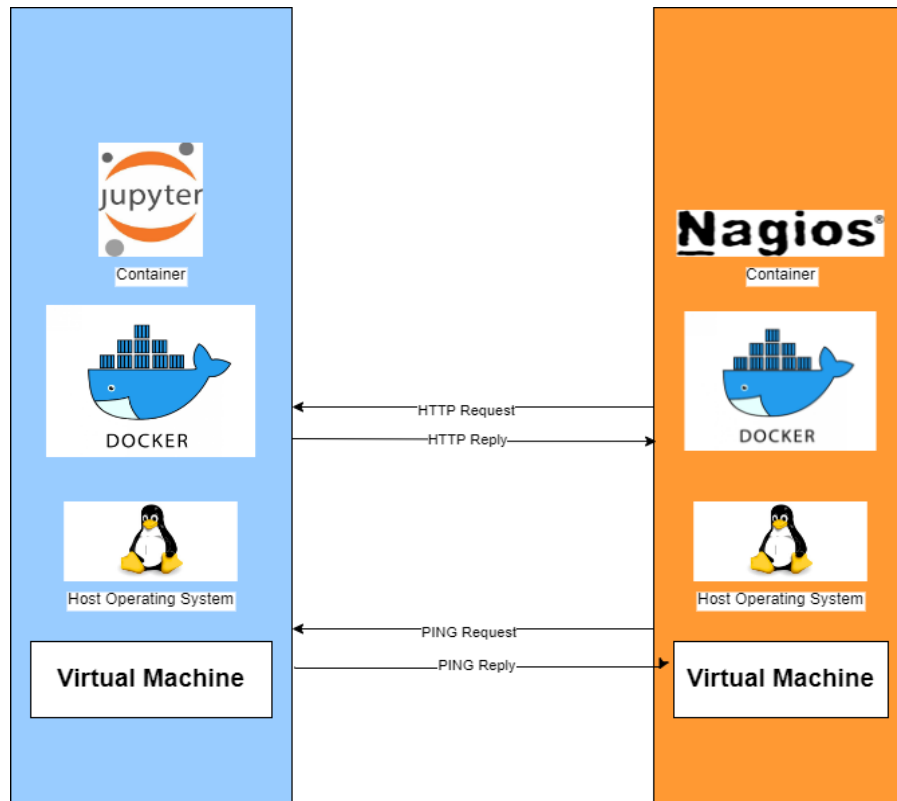
# Προτάσεις Βελτίωσης Καταχώρησης και Προτυποποίησης Παραπόνων

- ▶ Από τα παραπάνω δεδομένα και την ανάλυση που έχει γίνει από την ομάδα μας, προκύπτει ότι υπάρχουν αρκετές κατηγορίες παραπόνων όπου επικαλύπτονται ή έχουν λίγα παράπονα και θα μπορούσαν να συγχωνευθούν χωρίς να υπάρχει αυτή η κατάτμησή τους
- ▶ Επίσης φαίνεται να μην υπάρχει προτυποποιημένος τρόπος καταχώρησης τους ακόμα και για το μέσο στο οποίο και αναφέρονται.
- ▶ Από την πλευρά της ομάδας προτείνονται συγκεκριμένες κατηγορίες που θα βοηθήσουν και στην καλύτερη κατηγοριοποίηση των παραπόνων και την δυνατότητα κλιμάκωσης του μοντέλου που έχει αναπτυχθεί από την ομάδα.





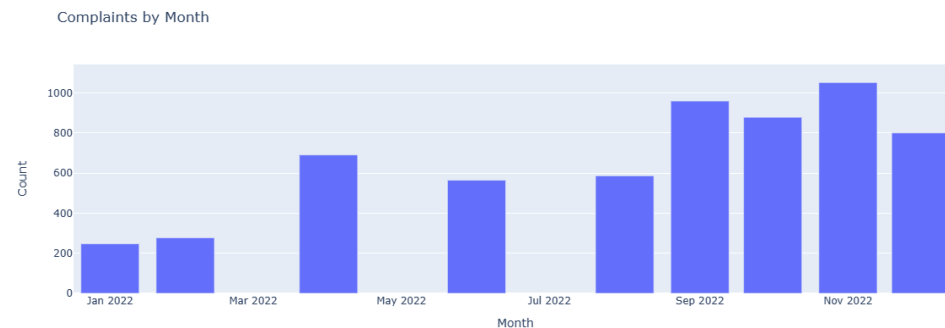
# Αρχιτεκτονική



- Όλα τα services μπορούν να «στηθούν» ως containers
- Το Nagios χρησιμοποιήθηκε ως ένα σύστημα ελέγχου της λειτουργίας της υποδομής (Virtual Machine) αλλά και του Service (Jupyter Notebook)

# Μηχανισμός Αυτόματης Ταξινόμησης Παραπόνων

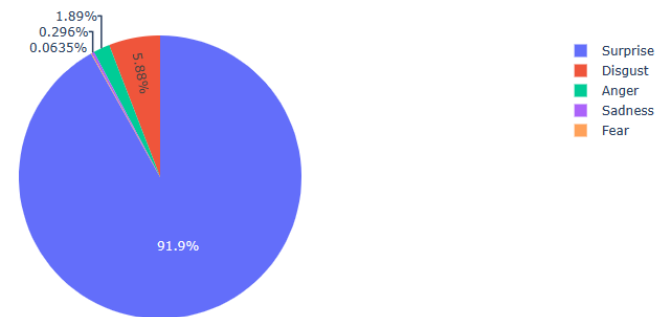
- ▶ Στα πλαίσια της εκπαίδευσης του μοντέλου αξιοποιήθηκαν 7100 παράπονα (αφαιρέθηκαν 38 καθώς ήταν μεγαλύτερα από 512 tokens που είναι και ο περιορισμός του μοντέλου)
- ▶ Χρησιμοποιήθηκε το νευρωνικό δίκτυο BERT με τη χρήση των προτεινόμενων κατηγοριών
- ▶ Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκε το 80% του συνόλου δεδομένων και το 20% για την εξέταση της αποτελεσματικότητάς του
- ▶ Η ακρίβεια του μοντέλου είναι αυτή τη στιγμή στο ~70%
- ▶ Παράλληλα έχει υλοποιηθεί και μηχανισμός επεξηγησιμότητας (explainability) του μοντέλου



# Ανάλυση Συναισθημάτων (Sentiment Analysis)

- ▶ Για την έυρεση των συναισθημάτων των επιβατών σχετικά με τα παράπονα χρησιμοποιήθηκε το παρακάτω λεξικό με κάποιες παραμετροποιήσεις:
  - ▶ [MKLab-ITI/greek-sentiment-lexicon: A lexicon to be used for sentiment analysis in Greek.](#)
  - ▶ Το λεξικό αυτό διαθέτει περίπου 2700 λέξεις και εξετάζει τα εξής συναισθήματα: αγανάκτηση, χαρά, φόβος, λύπη, έκπληξη, αποστροφή
- ▶ Για κάθε παράπονο υπάρχει λεπτομερής εικόνα για το score που έχει στα συγκεκριμένα συναισθήματα

Πρώτο Συναίσθημα



# Μηχανισμός Εύρεσης Επείγοντων Παραπόνων

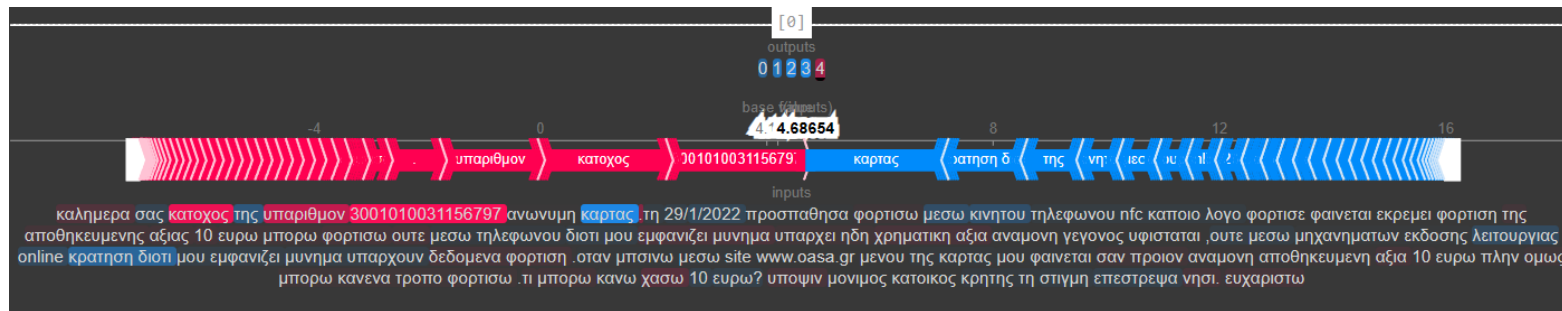
- ▶ Ο μηχανισμός εξετάζει αν τα παράπονα έχουν συγκεκριμένες «οντότητες» που δείχνουν παράπονα που ο επιβάτης έχει τραυματιστεί ή/και θέλει να προχωρήσει νομικά:
- ▶ Ο μηχανισμός έχει αναπτυχθεί με την χρήση συγκεκριμένων λέξεων που υποδηλώνουν τέτοιες ενέργειες
- ▶ Εδώ βασιζόμαστε σε ένα μοντέλο NER με τη χρήση του spaCY framework

## Επείγοντα Παράπονα

Unnamed: 0	init_text	SOS	ents
0	Αγαπητοί κύριοι καλησπέρα σας σήμερα 31/1 2022 και ώρα 7:30 το απόγευμα η κόρη μου ετών 12.5 βρισκόταν στο τέarma του 122 στην Σαρωνίδα με προορισμό το Λαγονήσι το δρομολόγιο των 7:40 δεν εκτελέστηκε ποτέ στη θέση του βρισκόταν το υπ αριθμόν ΧΕΗ 8394 λεωφορείο ΣΤΑΘΜΕΥΜΕΝΟ ΕΜΠΡΟΣΘΕΝ ΤΗΣ ΤΑΞΗΣ ΠΡΑΓΜΑΤΟΠΟΙΩΝΤΑΣ ΚΑΤΑ ΤΑ ΛΕΓΟΜΕΝΑ ΤΟΥ ΟΔΗΓΟΥ ΔΙΑΛΕΙΜΜΑ ΟΔΗΓΟΣ ΑΓΕΝΕΣΤΑΤΟΣ ΟΤΑΝ ΡΩΤΗΘΗΚΕ ΑΠΟ ΤΗΝ ΚΟΡΗ ΜΟΥ ΓΙΑ ΤΟ ΤΙ ΩΡΑ ΘΑ ΦΥΓΕΙ ΑΡΚΕΣΤΗΚΕ ΣΤΟ ΝΑ ΠΕΙ ΣΕ 45 ΛΕΠΤΑ ΈΧΟΝΤΑΣ ΚΛΕΙΣΤΕΣ ΤΙΣ ΠΟΡΤΕΣ ΚΑΙ ΑΦΗΝΟΝΤΑΣ ΤΟΝ ΚΟΣΜΟ ΣΤΟ ΚΡΥΟ ΕΝΩ Ο ΟΔΗΓΟΣ ΒΡΙΣΚΟΤΑΝ ΜΕ ΆΛΛΑ ΔΥΟ ΑΤΟΜΑ ΕΝΤΟΣ ΤΟΥ ΛΕΩΦΟΡΕΙΟΥ ΚΑΙ ΜΙΛΑΓΕ ΕΙΜΑΙ ΟΔΗΓΟΣ ΤΑΞΙ ΠΡΑΓΜΑΤΟΠΟΙΗΣΑ ΤΟ ΔΡΟΜΟΛΟΓΙΟ ΒΑΡΚΙΖΑ ΣΑΡΩΝΙΔΑ ΧΩΡΙΣ ΤΟ ΔΡΟΜΟ ΜΟΥ ΝΑ ΣΥΝΑΝΤΗΣΩ ΤΟ 122 ΕΝΩ ΑΡΚΕΤΑ ΛΕΠΤΑ ΑΡΓΟΤΕΡΑ ΤΟ ΥΠ ΑΡΙΘΜΟΝ ΧΕΗ8390 ΠΡΑΓΜΑΤΟΠΟΙΗΣΕ ΚΑΝΟΝΙΚΑ ΤΟ ΔΡΟΜΟΛΟΓΙΟ ΤΟΥ ΣΤΙΣ 8:05 ΜΕ ΕΥΓΕΝΕΣΤΑΤΟ ΟΔΗΓΟ ΑΓΑΠΗΤΟΙ ΚΥΡΙΟΙ ΤΟΥ ΟΑΣΑ ΚΑΙ ΤΗΣ ΌΣΟΙ ΕΙΝΑΙ ΠΟΛΙΤΙΚΗ ΣΑΣ ΝΑ ΑΦΗΝΕΤΕ ΠΑΙΔΙΑ ΜΕΣΑ ΣΤΟ ΚΡΥΟ ΑΚΟΜΑ ΚΑΙ ΑΝ ΤΟ ΛΕΩΦΟΡΕΙΟ ΚΑΝΕΙ ΔΙΑΛΕΙΜΜΑ Η ΚΟΡΗ ΜΟΥ ΠΟΥ ΈΧΑΣΕ ΤΟ ΜΑΘΗΜΑ ΣΤΟ ΛΑΓΟΝΗΣΙ ΚΑΙ ΠΟΥ ΕΓΩ ΈΤΡΕΞΑ ΝΑ ΠΡΟΛΑΒΩ ΘΑ ΜΕ ΑΠΟΖΗΜΙΩΣΕΤΕ ΈΧΕΤΕ ΣΚΟΠΟ ΝΑ ΦΤΙΑΞΕΤΕ ΤΗ ΓΡΑΜΜΗ 122 ΠΟΤΕ ΜΕ ΜΕΓΑΛΗ ΑΓΑΝΑΚΤΗΣΗ ΜΑΓΚΑΣ ΚΩΝΣΤΑΝΤΙΝΟΣ ΣΤΗ ΔΙΑΘΕΣΗ ΣΑΣ ΦΩΤΟΓΡΑΦΙΑ ΜΕ ΤΟ ΛΕΩΦΟΡΕΙΟ ΝΑ ΚΑΝΕΙ ΔΙΑΛΕΙΜΜΑ ΜΠΡΟΣΤΑ ΣΤΗ ΣΤΑΣΗ ΤΟΥ ΤΕΡΜΑΤΟΣ ΚΑΤΙ ΠΟΥ ΚΑΤΑ ΤΗ ΓΝΩΜΗ ΜΟΥ ΕΙΝΑΙ ΑΠΑΡΑΔΕΚΤΗ	1	(αποζημιώσετε,)
1	Παρακαλώ πολύ, σήμερα, Κυριακή, 30 /01 / 2022 και ώρα 7:32 μ.μ. το λεωφορείο με αριθμό δρομολογίου 509, με κατεύθυνση από Ζηρίνιο προς Άγιο Στέφανο Αττικής, δεν σταμάτησε στην στάση επί της λεωφόρου Θησέως, στην Εκάλη, στην πλατεία Βασιλέως Παύλου, παρόλο που του έκανε σήμα η κόρη μου, η οποία περίμενε μέσα στη νύχτα και στην παγωνιά, Αντιθέτως ο οδηγός έτρεχε με υπερβολική, παράνομη ταχύτητα και δεν σταμάτησε Παρακαλώ πολύ, όπως με ενημερώσετε για τα στοιχεία του συγκεκριμένου οδηγού, για το συγκεκριμένο δρομολόγιο, γιατί θα ασκήσω μήνυση εναντίον του, ασκώντας όλα τα νόμιμα δικαιώματά μου, επιφυλασσόμενη αυτών, καθόσον είμαι δικηγόρος. Σε τόσο δύσκολους καιρούς που ζούμε και ενόψει τόσο δύσκολων καιρικών συνθηκών , τέτοιες συμπεριφορές είναι απαράδεκτες και αδιανόητες !!! Ευχαριστώ πολύ ΕΛΕΝΗ ΠΑΝΤΑΖΗ	1	(μηνυση,)

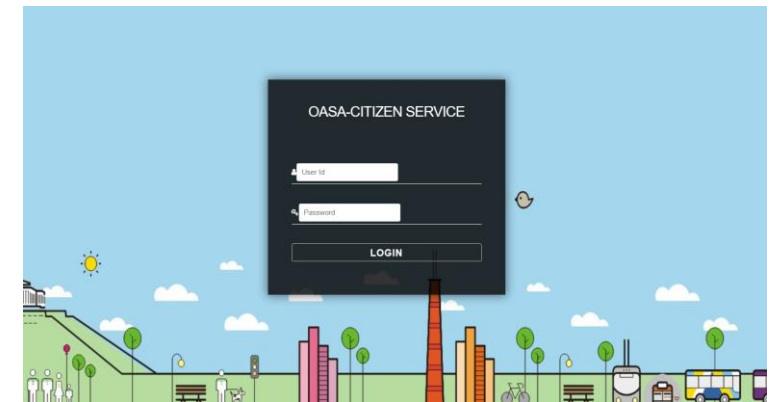
# Επεξηγησιμότητα (Explainability)

- ▶ Μηχανισμός επεξηγησιμότητας (ως proof of concept) έχει υλοποιηθεί για τον αυτόματο ταξινομητή, ώστε να μπορεί να επεξηγηθεί η λειτουργία του συγκεκριμένου μηχανισμού
- ▶ Έχει χρησιμοποιηθεί η μέθοδος SHAP (SHapley Additive exPlanations), όπου μπορεί να αξιοποιηθεί σε κάθε μοντέλο μηχανικής μάθησης.



# Οπτικοποίηση

- ▶ Για την οπτικοποίηση των δεδομένων έχει δημιουργηθεί ένα σχετικό dashboard, που εμφανίζει το σύνολο των πληροφοριών των υποσυστημάτων που έχουν δημιουργηθεί
- ▶ Χρησιμοποιήθηκε η βιβλιοθήκη ανοιχτού κώδικα dash plotly για τη δημιουργία του περιβάλλοντος και των γραφημάτων



# Επόμενα Βήματα

- ▶ Βελτιστοποίηση Μηχανισμών Αυτόματης Κατηγοριοποίησης Παραπόνων και εύρεσης παραπόνων που χρήζουν άμεσης βελτίωσης
- ▶ Ενσωμάτωση και άλλων γραφημάτων σχετικών με την ανάλυση των παραπόνων
- ▶ Χρήση των επείγοντων παραπόνων για την εκπαίδευση του υφιστάμενου chatbot του οργανισμού
- ▶ Αξιοποίηση Large Language Models (LLMs) όπως το ChatGPT στην βελτιστοποίηση των μηχανισμών που έχουν υλοποιηθεί

Thank  
You!