



ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΤΗΛΕΜΑΤΙΚΗΣ

2η Εργασία: Επικοινωνία ανθρώπου-μηχανής

Ομάδα:

Καζάκος Χρήστος, it22033

Κωνσταντίνος Κατσάρας, it22045

Μανούσος Λιναρδάκης, it22064

Παρατήρηση: Για να κάνουμε τις προβλέψεις, κανονικοποιούμε πρώτα τα δεδομένα μας (καλύτερα αποτελέσματα).

Περιγραφή 1

Αρχικά, γίνεται εισαγωγή των απαιτούμενων βιβλιοθηκών. Στη συνέχεια, γίνεται αντιγραφή του συνόλου δεδομένων από το Colab. Έπειτα, πραγματοποιείται η φόρτωση του συνόλου δεδομένων με εκτύπωση στατιστικών για τιμές *nan*, αριθμητικές και κατηγορικές τιμές. Αφαιρείται η στήλη 'type', καθώς αυτή δεν περιλαμβάνεται στις αναφερόμενες στήλες της εκφώνησης (περιγραφή 1). Στη συνέχεια, αφαιρούνται οι τιμές *nan*, διότι αποτελούν έναν εξαιρετικά μικρό ποσοστό στο σύνολο δεδομένων. Τέλος, προχωράμε στην κανονικοποίηση των δεδομένων με σκοπό την επίτευξη βέλτιστων αποτελεσμάτων.

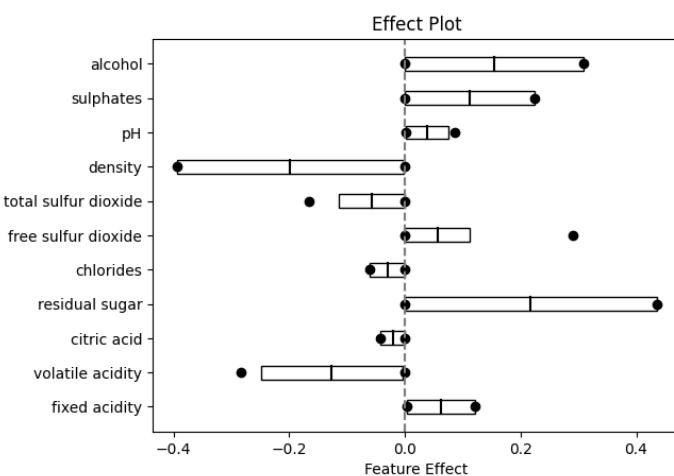
Ζητούμενα

Ζητούμενο Α:

Το R-Squared βγάζει **0.35**, κάτι που δεν είναι τόσο καλό στη συνολική ικανότητα του μοντέλου να εξηγεί τα δεδομένα. Το σύνηθες εύρος τιμών του R-squared είναι μεταξύ '0' (το μοντέλο δεν εξηγεί καθόλου τα δεδομένα) και '1' (μοντέλα που εξηγούν το σύνολο της διακύμανσης στα δεδομένα).

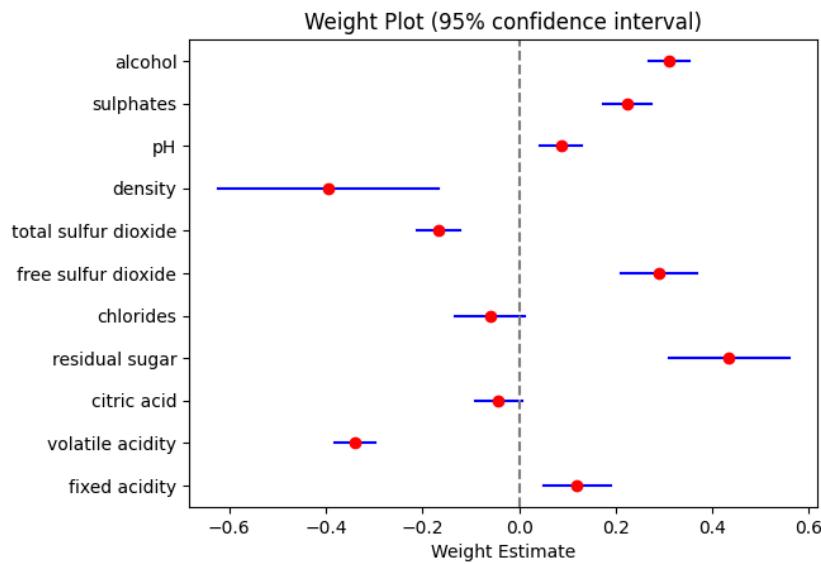
Ζητούμενο Β:

Effect Plot:



Σύμφωνα με το παραπάνω διάγραμμα, τα χαρακτηριστικά με τη μεγαλύτερη θετική συνεισφορά είναι το "density" όπως μπορούμε να δούμε από το μέγεθος των κουτιών (ύστερα ακολουθεί το γνώρισμα "alcohol"). Τη μεγαλύτερη αρνητική συνεισφορά την έχει "residual sugar" και έπειτα ακολουθεί το "volatile acidity".

Weight Plot:

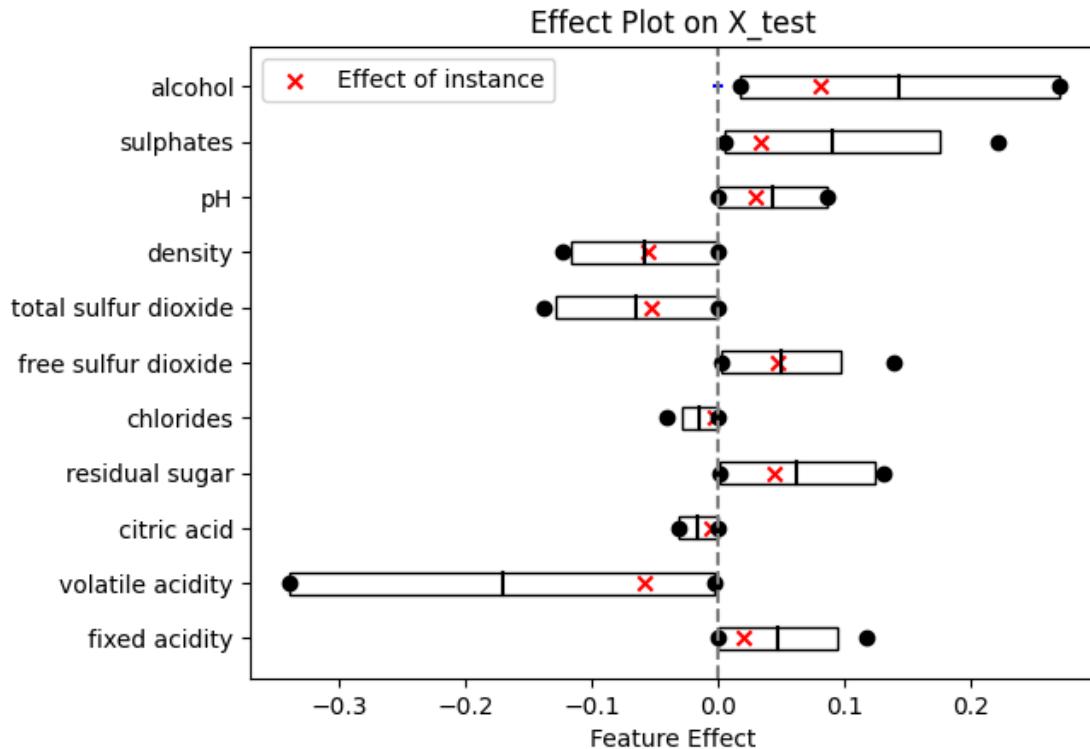


Παρατηρούμε ότι η επίδραση του γνωρίσματος "residual sugar" έχει τη μεγαλύτερη θετική επίδραση (σε σχέση με τα υπόλοιπα γνωρίσματα) και μετά ακολουθεί το free sulfur dioxide. Την μεγαλύτερη αρνητική επίδραση την έχει το γνώρισμα "density".

Από την άλλη τα γνωρίσματα όπως το pH, chlorides, citric acid, fixed acidity δεν έχουν στατιστικά σημαντική επίδραση στη πρόβλεψη.

Παρατήρηση: Οι προβλέψεις έχουν διάστημα εμπιστοσύνης 95%.

Ζητούμενο C:



Παρατηρούμε ότι το predicted value του συγκεκριμένου δείγματος είναι **μεγαλύτερο** από το μέσο predicted value.

Από το σχήμα, μπορούμε να δούμε ότι η επίδραση του instance για κάθε γνώρισμα είναι εντός του αναμενόμενου range. Συγκεκριμένα, οι επιδράσεις του δείγματος φαίνονται από τα κόκκινα "X" στο διάγραμμα.

Περιγραφή 2

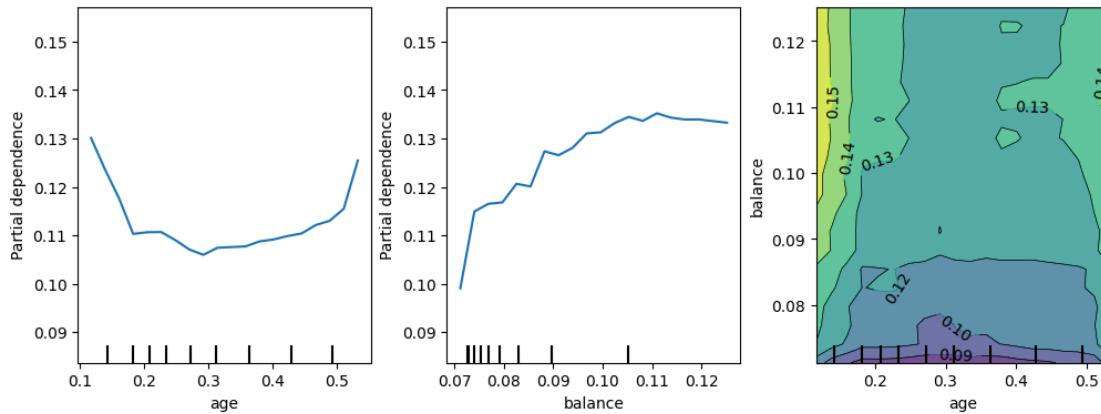
Κώδικας

Αρχικά, πραγματοποιούμε την εισαγωγή των απαιτούμενων βιβλιοθηκών. Στη συνέχεια, φορτώνουμε το σύνολο δεδομένων, όπως περιγράφεται στο εγχειρίδιο του συνόλου δεδομένων ([import in python](#)). Στο στάδιο της προεπεξεργασίας, μετατρέπουμε τις κατηγορικές μεταβλητές σε αριθμητικές μέσω κωδικοποίησης one-hot. Επιπλέον, προβαίνουμε στην κανονικοποίηση των τιμών, ώστε να κινούνται στο εύρος από 0 έως 1. Τέλος, εκτελούμε τον κώδικα που αναφέρεται στη δεύτερη περιγραφή..

Ζητούμενα

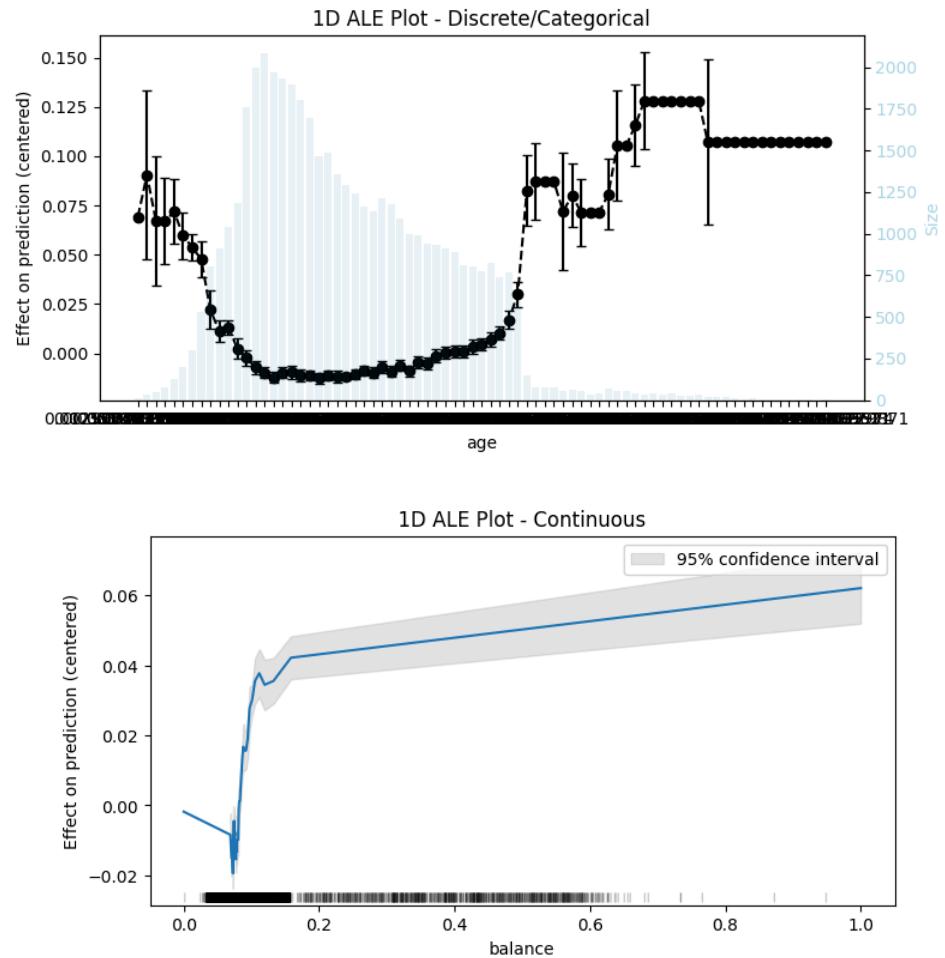
Ζητούμενο Α

1-way vs 2-way of numerical PDP

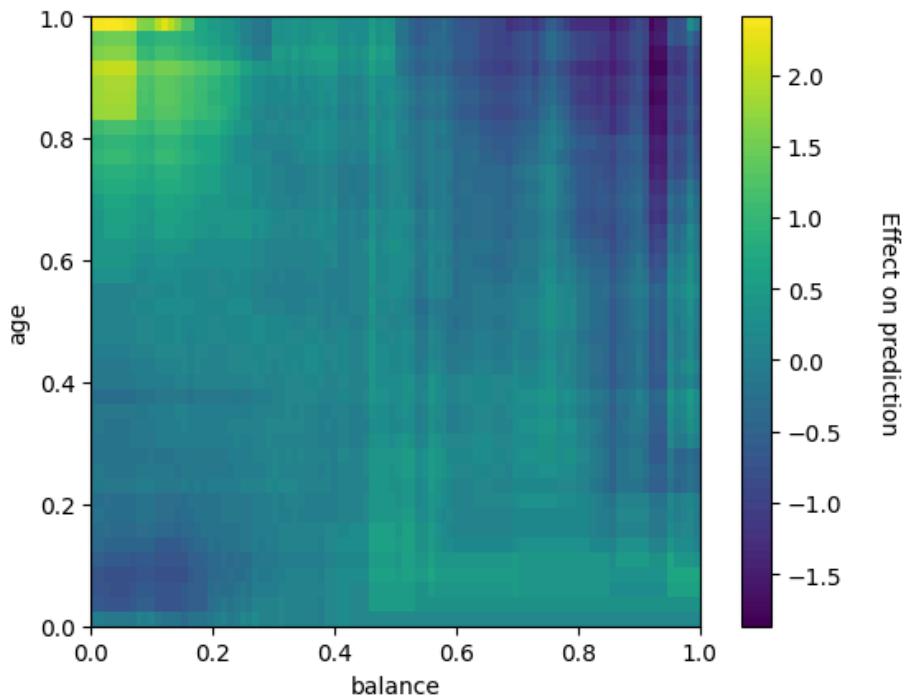


- Το πρώτο (αριστερά) γράφημα δείχνει τη μερική εξάρτηση έναντι της 'Ηλικίας'. Μια μπλε γραμμή σε αυτό το πρώτο διάγραμμα υποδεικνύει μια αρχική μείωση της μερικής εξάρτησης καθώς αυξάνεται η ηλικία, ακολουθούμενη από μια αύξηση σχηματίζοντας σχήμα U.
- Το μεσαίο γράφημα απεικονίζει μερική εξάρτηση έναντι του 'Μέσου Υπόλοιπου', με τιμές που κυμαίνονται από περίπου 0.07 έως περίπου 0.12 για την ισορροπία. Βλέπουμε ότι όσο αυξάνεται το balance, αυξάνεται και η μερική εξάρτηση, μέχρι που "σταθεροποιείται" στο 0.1.
- Το τελευταίο δεξιά γράφημα δείχνει πως επιδρά το μέσο υπόλοιπο με την ηλικία στην μερική εξάρτηση. Παρατηρούμε ότι αν το $age < 0.15$ (περίπου) και το balance είναι στο διάστημα $[0.1, 0.12]$ η μερική εξάρτηση είναι ψηλότερη (άρα η επίδραση στην πρόβλεψη του μοντέλου είναι μεγαλύτερη).

Ζητούμενο Β

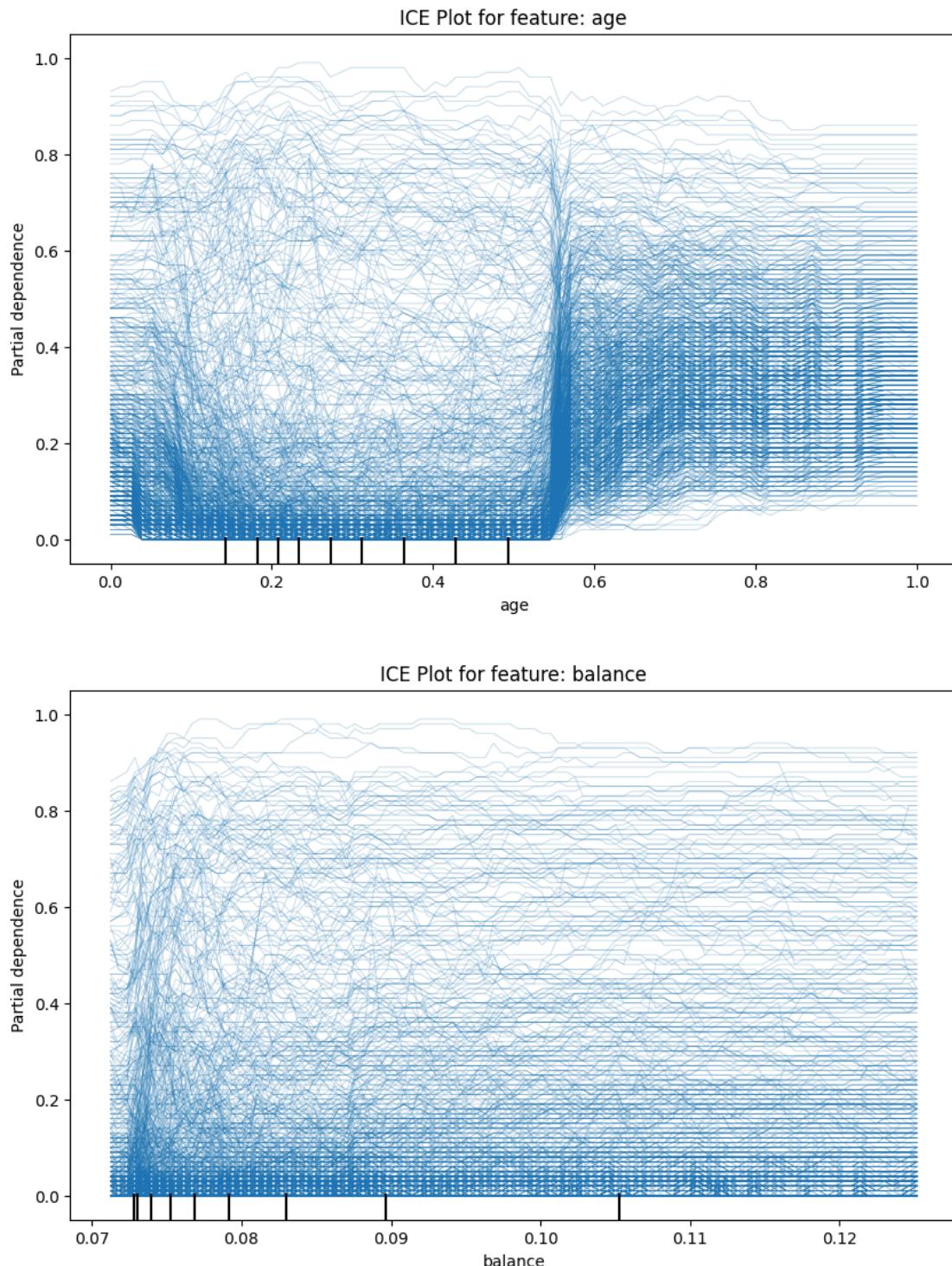


2D ALE Plot



- Στο πρώτο γράφημα, παρατηρούμε ότι αν το age βρίσκεται (περίπου) στο [0.8, 1] έχει υψηλή επίδραση στις προβλέψεις.
- Στο δεύτερο γράφημα, βλέπουμε ότι όσο αυξάνεται το balance, τόσο μεγαλύτερη επίδραση έχει στις προβλέψεις (γενικώς όμως έχει χαμηλότερη επίδραση από το 'age').
- Συνδυαστικά το balance, prediction στο τελευταίο γράφημα παρατηρούμε ότι το μεγαλύτερο effect στο prediction το έχουμε στις τιμές [0, 0.2] για το balance και [0.8, 1] για το age.

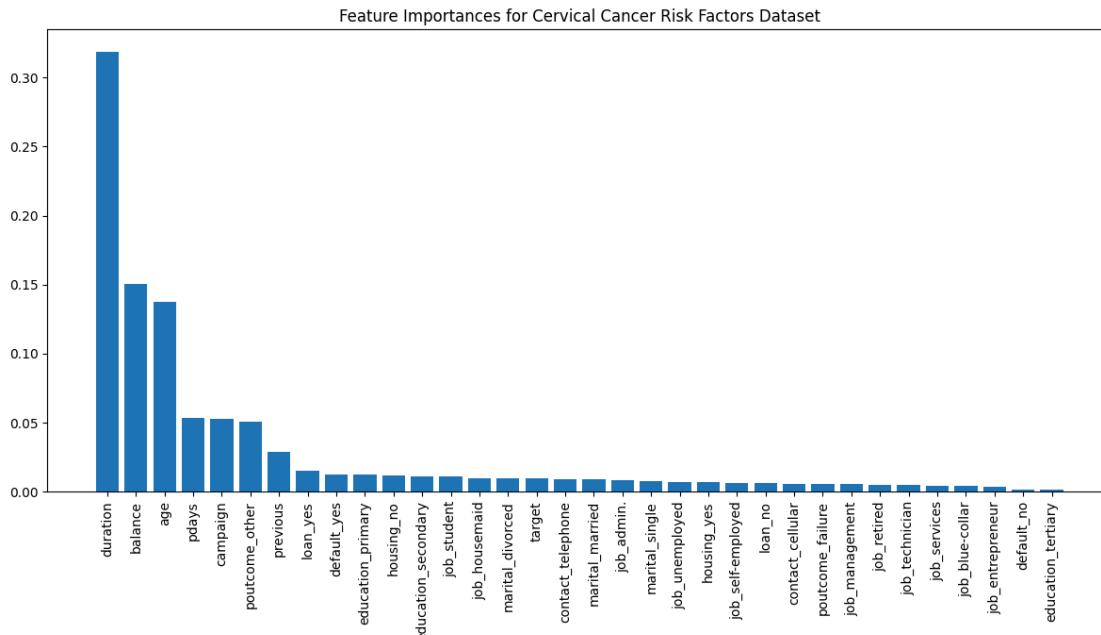
Ζητούμενο C



- Τα γραφήματα ICE αναπαριστούν μία γραμμή ανά εμφάνιση που δείχνει πώς μεταβάλλεται η πρόβλεψη όταν αλλάζει ένα χαρακτηριστικό.

- Στο πρώτο διάγραμμα παρατηρούμε μια γενική άνοδο της συσχέτισης της ηλικίας όταν εκείνη ξεπερνά το 0.55, ενώ όταν είναι μικρότερη από 0.55 είναι συγκριτικά χαμηλή.
- Στο δεύτερο διάγραμμα παρατηρούμε πως το balance έχει σταθερά θετική και ελαφρώς αυξανόμενη συσχέτιση με την προβλεπόμενη τιμή καθώς δε συναντάμε κάποια απότομη μεταβολή του Dependence κατά την μεταβολή της τιμής του balance.

Ζητούμενο D



Παρατηρούμε ότι η διάρκεια είναι το πιο σημαντικό feature (σήμφωνα με τη μετρική σπουδαιότητας) και μετά ακολουθούν το μέσο υπόλοιπο και η ηλικία. Τα υπόλοιπα feautures έχουν σχετικά χαμηλή σημαντικότητα.

Ζητούμενο E:

Από τα ερωτήματα A, B, C βλέπουμε συνολικά τις συσχετίσεις μεταξύ των age, balance καθώς και την επίδραση που έχουν προς τις προβλέψεις του μοντέλου.

Από τα διαγράμματα μπορούμε να βγάλουμε τα εξής συμπεράσματα / παρατηρήσεις:

1. Το balance και το age έχουν "αντίστροφη" σχέση στα διαγράμματα του A και B. Σγκεκρικά, στο A η μερική εξάρτηση είναι μεγαλύτερη όταν το balance είναι μεγάλο και το age μικρό. Αντίθετα, στο B υπάρχει μεγαλύτερη επίδραση στη πρόβλεψη όταν το age είναι μεγάλο και το balance μικρό.
2. Το διάγραμμα C "συνδυάζει" τα αποτέλεσμα του A και B, καθώς σε αυτό όσο "μεγαλώνει" το age, αυξάνεται η μερική εξάρτηση (όπως το B) και όσο μεγαλώνει το balance, αυξάνεται επίσης η μερική εξάρτηση (όπως το A).

Περιγραφή 3

Κώδικας

Αρχικά, προχωρούμε στην κανονικοποίηση των τιμών στο εύρος από 0 έως 1 για βέλτιστα αποτελέσματα. Στη συνέχεια, αντιγράφουμε το σύνολο δεδομένων από το Colab, φορτώνουμε το σύνολο δεδομένων και πραγματοποιούμε την κανονικοποίηση των τιμών στο εύρος 0 έως 1 για βελτιστοποίηση των αποτελεσμάτων. Έπειτα, εκτελούμε τον κώδικα που αναφέρεται στην τρίτη περιγραφή.

Ζητούμενα

Ζητούμενο A

```
Surrogate Decision Tree Accuracy: 69.51%
|--- rm <= 0.63
|   |--- lstat <= 0.35
|   |   |--- lstat <= 0.09
|   |   |   |--- ptratio <= 0.76
|   |   |   |   |--- rm <= 0.60
|   |   |   |   |   |--- value: [0.49]
|   |   |   |   |   |--- rm >  0.60
|   |   |   |   |   |--- value: [0.58]
|   |   |   |--- ptratio >  0.76
|   |   |   |--- lstat <= 0.05
|   |   |   |   |--- value: [0.85]
|   |   |   |--- lstat >  0.05
|   |   |   |   |--- value: [0.91]
|--- lstat >  0.09
|   |--- lstat <= 0.22
|   |   |--- rm <= 0.49
|   |   |   |--- value: [0.37]
|   |   |--- rm >  0.49
|   |   |   |--- value: [0.45]
|--- lstat >  0.22
|   |--- indus <= 0.09
|   |   |--- value: [0.46]
|   |--- indus >  0.09
|   |   |   |--- value: [0.35]
|--- lstat >  0.35
|   |--- crim <= 0.06
|   |   |--- nox <= 0.30
```

```
| | | | | --- nox <= 0.18
| | | | | --- value: [0.28]
| | | | --- nox >  0.18
| | | | | --- value: [0.35]
| | | --- nox >  0.30
| | | | --- lstat <= 0.45
| | | | | --- value: [0.28]
| | | | --- lstat >  0.45
| | | | | --- value: [0.22]
| --- crim >  0.06
| | --- lstat <= 0.51
| | | --- nox <= 0.41
| | | | --- value: [0.27]
| | | --- nox >  0.41
| | | | --- value: [0.20]
| | --- lstat >  0.51
| | | --- nox <= 0.60
| | | | --- value: [0.18]
| | | --- nox >  0.60
| | | | --- value: [0.10]
--- rm >  0.63
| --- rm <= 0.74
| | --- nox <= 0.56
| | | --- lstat <= 0.10
| | | | --- nox <= 0.48
| | | | | --- value: [0.65]
| | | | --- nox >  0.48
| | | | | --- value: [0.89]
| | --- lstat >  0.10
| | | --- rm <= 0.67
| | | | --- value: [0.50]
| | | --- rm >  0.67
| | | | --- value: [0.63]
| --- nox >  0.56
| | --- rm <= 0.69
| | | --- value: [0.20]
| | --- rm >  0.69
| | | | --- dis <= 0.07
| | | | | --- value: [0.35]
```

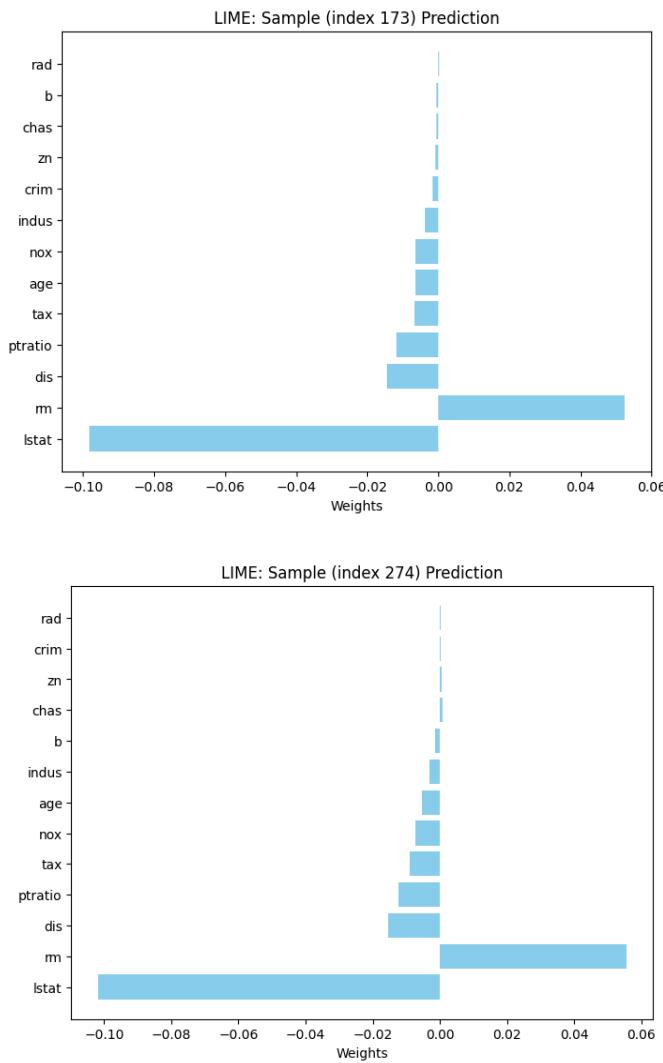
```
| | | | | --- dis >  0.07
| | | | | --- value: [0.29]
| --- rm >  0.74
| --- ptratio <= 0.61
| | --- ptratio <= 0.30
| | | --- rm <= 0.82
| | | | --- value: [0.89]
| | | --- rm >  0.82
| | | | --- value: [0.98]
| | --- ptratio >  0.30
| | | --- age <= 0.62
| | | | --- value: [0.88]
| | | --- age >  0.62
| | | | --- value: [0.82]
| --- ptratio >  0.61
| | --- indus <= 0.44
| | | --- value: [0.74]
| | --- indus >  0.44
| | | --- value: [0.51]
```

Παρατηρούμε πως το χαρακτηριστικό RM (Μέσος αριθμός δωματίων ανά κατοικία) είναι από τα βασικά χαρακτηριστικά και καθορίζει σημαντικά τις προβλέψεις αξίας του ακινήτου, καθώς το συναντάμε σε πολλά σημεία κόμβων. Συγκεκριμένα, χαμηλές τιμές του 'RM' (<0.63) φαίνεται να συσχετίζονται με χαμηλές τιμές του στόχου (τιμή κατοικίας), ενώ υψηλότερες τιμές του 'RM' (> 0.74) συσχετίζονται με υψηλότερες τιμές κατοικιών.

Το χαρακτηριστικό LSTAT' (ποσοστό κατοίκων χαμηλού εισοδήματος) είναι επίσης αρκετά σημαντικός παράγοντας και παρατηρούμε ότι χαμηλά ποσοστά συνδέονται με υψηλότερες προβλεπόμενες τιμές κατοικιών, ενώ υψηλά ποσοστά (> 0.35) συνδέονται με χαμηλότερες.

Τέλος, το χαρακτηριστικό CRIM (αναλογία εγκληματικότητας) φαίνεται να έχει και αυτό κάποια σημασία καθώς χαμηλές τιμές συνδέονται με μεγαλύτερες προβλεπόμενες τιμές.

Ζητούμενο Β



Για τα 2 επιλεγμένα δείγματα παρατηρούμε πως το χαρακτηριστικό LSTAT έχει αρνητική συσχέτιση με τη τιμή των κατοικιών. Συνεπώς όσο αυξάνεται η τιμή του LSTAT, τόσο χαμηλή τιμή περιμενουμε να είναι η τιμή της κατοικίας και ομοίως όσο μειώνεται η τιμή του LSTAT περιμένουμε τη τιμή της κατοικίας να είναι υψηλή.

Όσον αφορά το χαρακτηριστικό RM παρατηρούμε πως έχει θετική συσχέτιση με την τιμή της κατοικίας. Δηλαδή, όταν η τιμή του RM είναι υψηλή περιμένουμε να είναι υψηλή και η τιμή της κατοικίας και όταν η τιμή του RM είναι χαμηλή περιμένουμε τη τιμή της κατοικίας να είναι εξίσου χαμηλή.

Τέλος, το χαρακτηριστικό CRIM βλέπουμε πως έχει μια ελαφρά αρνητική συσχέτιση αλλά κατά κύριο λόγο δεν επηρεάζει σημαντικά τη τιμή μιας κατοικίας καθώς σύμφωνα με το διάγραμμα η συσχέτισή του είναι κοντά στο 0.

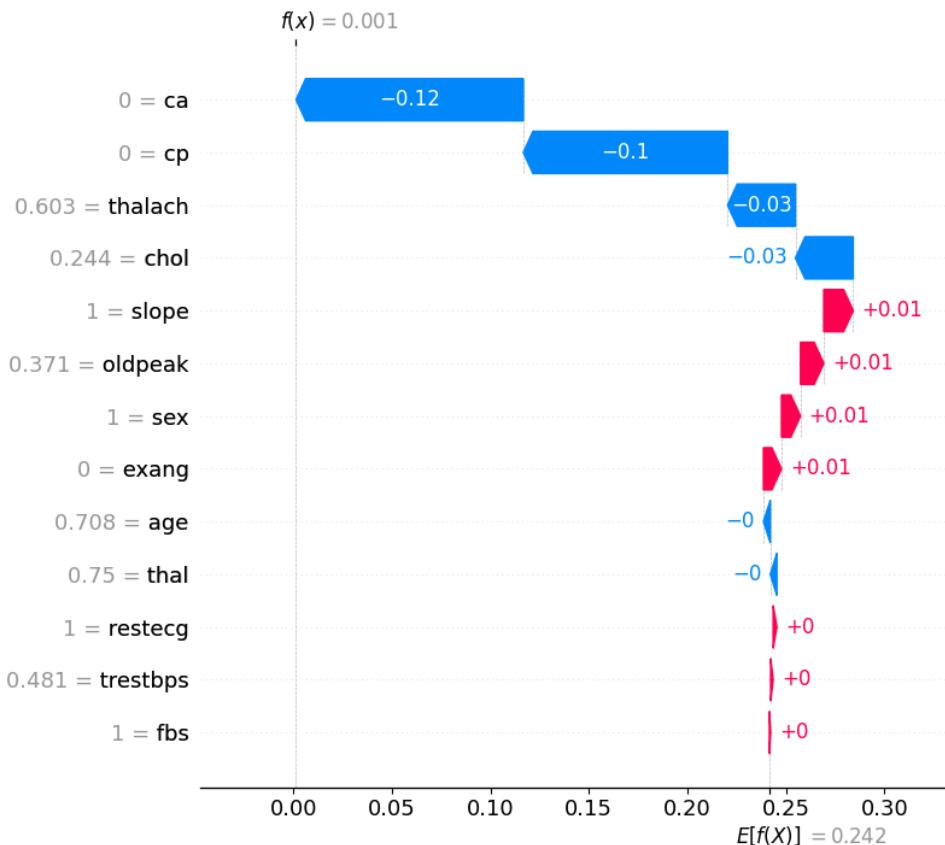
Περιγραφή 4

Κώδικας

Αρχικά, γίνεται η εισαγωγή των απαιτούμενων βιβλιοθηκών. Στη συνέχεια, προβαίνουμε στη φόρτωση του συνόλου δεδομένων, όπως περιγράφεται στο εγχειρίδιο του συνόλου δεδομένων ([import in python](#)). Στο στάδιο αυτό, αφαιρούμε τις τιμές `nan` και προχωρούμε στην κανονικοποίηση των δεδομένων. Στη συνέχεια, εκπαιδεύουμε και αρχικοποιούμε το SHAP (SHapley Additive exPlanations).

Ζητούμενα

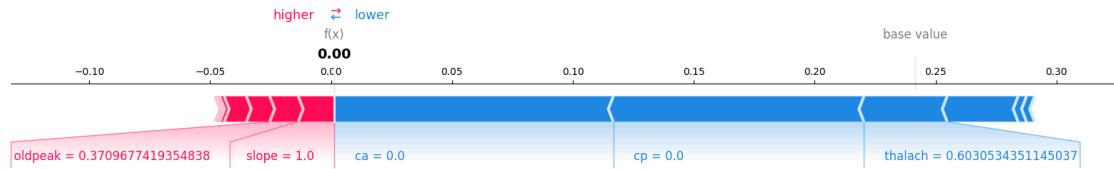
Ζητούμενο Α:



Παρατηρούμε ότι, **για το συγκεκριμένο δείγμα**, τα γνωρίσματα `ca`, `cp`, `thalach`, `chol`, `thal` και `age` (μπλέ μπάρες) έχουν αρνητική επίδραση στην πρόβλεψη, ενώ τα γνωρίσματα όπως το `slope`, `oldpeak`, `sex`, `exang`, `restecg`, `trestbps` και `fbs` (κόκκινες μπάρες) έχουν θετική επίδραση.

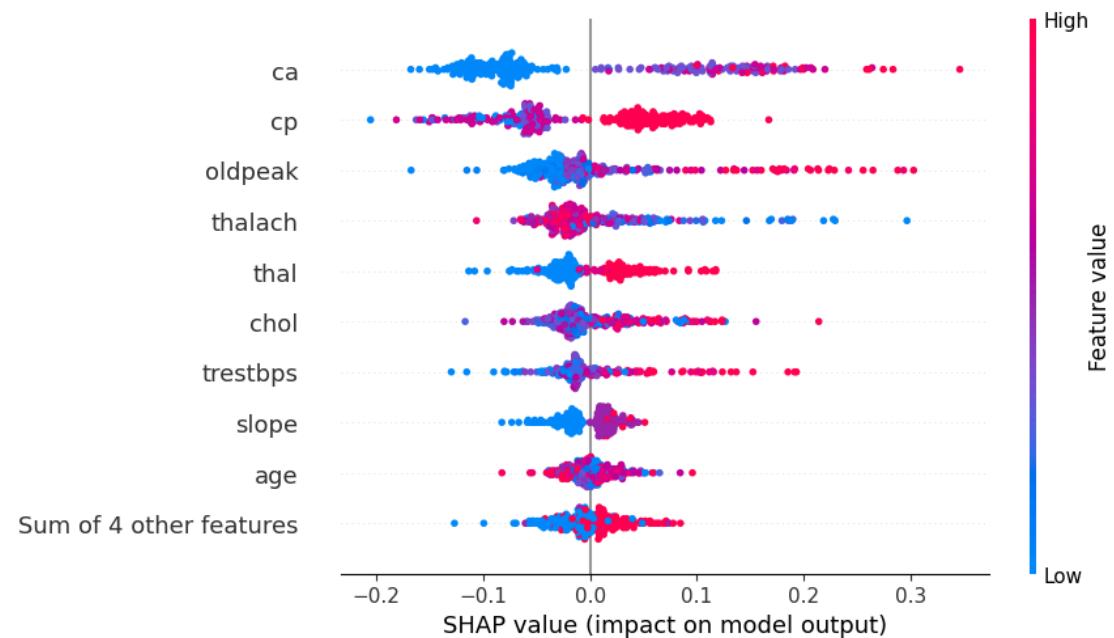
Επιπρόσθετα, το γνώρισμα `ca` έχει τη μεγαλύτερη αρνητική επίδραση (μετά ακολουθεί το `cp` και ύστερα το `thalach`), ενώ το γνώρισμα `slope` έχει τη μεγαλύτερη θετική επίδραση (και έπειτα ακολουθεί το `oldpeak`).

Ζητούμενο Β:



Παρατηρούμε ότι τα γνωρίσματα "oldpeak" και "slope" αυξάνουν την πρόβλεψη (σε σχέση με την μέση πρόβλεψη), ενώ τα features όπως το "ca", "cp" και "thalach" μειώνουν την πρόβλεψη **για το συγκεκριμένο δείγμα**.

Ζητούμενο Κ:

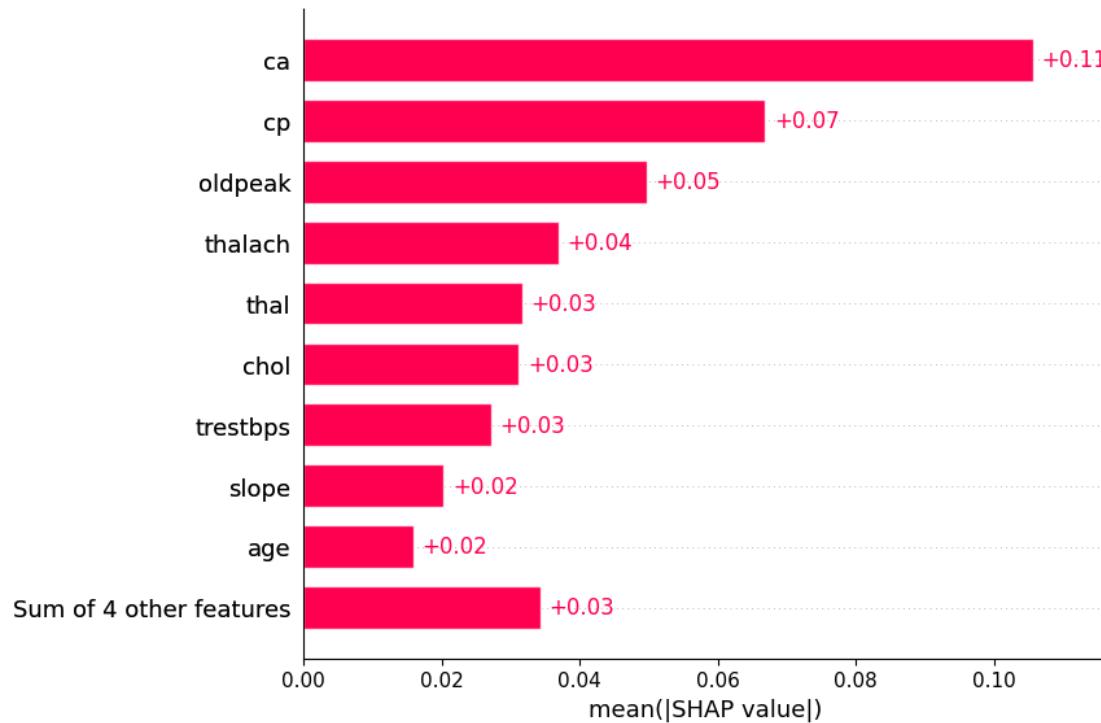


Παρατηρούμε ότι τα γνωρίσματα όπως το ca, cp, oldpeak, thalach έχουν μεγαλύτερη επίδραση στη πρόβλεψη.

Επίσης, βλέπουμε ότι χαμηλός αριθμός ca μειώνει το αποτέλεσμα της πρόβλεψης, ενώ όσο αυξάνεται το ca, αυξάνεται και το αποτέλεσμα (της πρόβλεψης). Όμοια συμπεριφορά έχουν και τα γνωρίσματα cp, oldpeak.

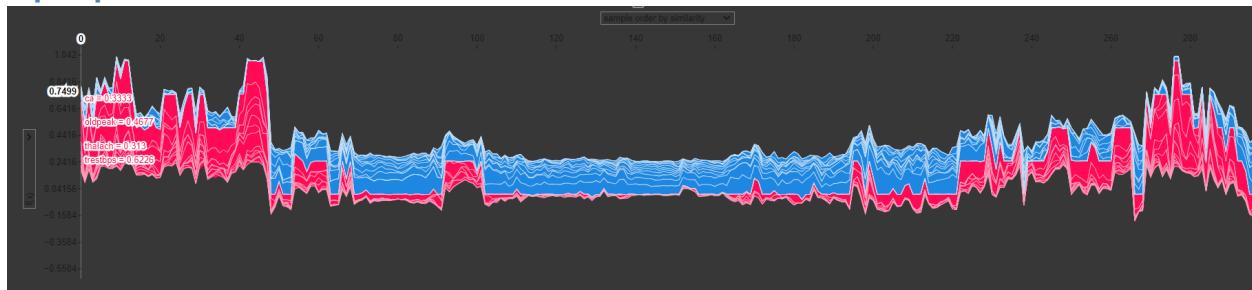
Από την άλλη πλευρά, το γνώρισμα thalach όσο μειώνεται, αυξάνει το αποτέλεσμα της πρόβλεψης, ενώ όσο αυξάνεται, το μειώνει (το αποτέλεσμα).

Ζητούμενο D:

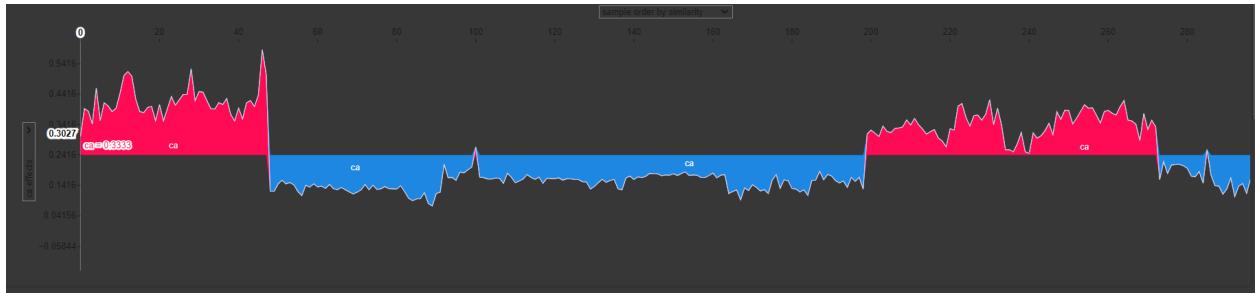


Παρατηρούμε ότι τα γνωρίσματα όπως το ca, cp έχουν τις μεγαλύτερες επιπτώσεις στη μέση τιμή πρόβλεψης (κατά 11% και 7% αντίστοιχα). Ύστερα, ακολουθεί το oldpeak με επίπτωση (στη μέση τιμή πρόβλεψης) 5%. Τα υπόλοιπα features έχουν επίπτωση μικρότερη ή ίση του 4%.

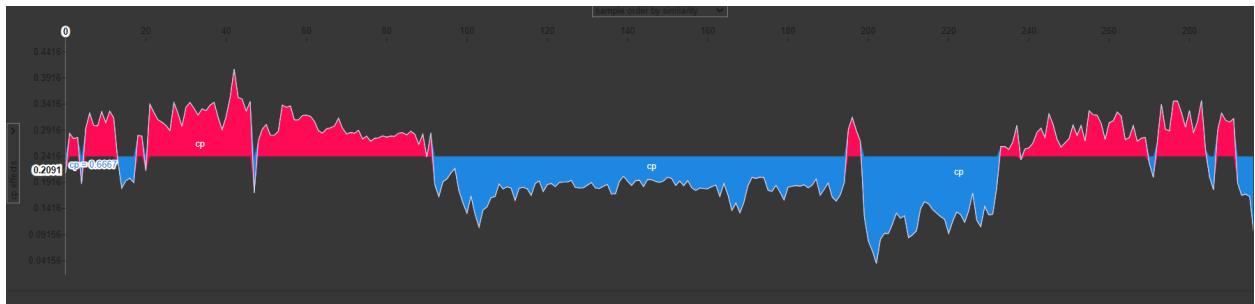
Ζητούμενο E:



$$y = f(x)$$



$y = ca$ effects



$y = cp$ effects

Παρατηρήσεις:

Βλέποντας το παραπάνω γράφημα (y axis = $f(x)$ και x axis = sample order by similarity) και συγκρίνοντας το με το y axis = ca effects, βλέπουμε ότι το ca έχει μεγάλη επίπτωση στη πρόβλεψη (αφού όταν αυξάνεται ή μειώνεται, πολλές φορές αλλάζει με όμοιο τρόπο και η πρόβλεψη).

Σχεδόν το ίδιο (με λιγότερη επίδραση από το ca) μπορούμε να παρατηρήσουμε και για το cp (αν y axis = cp effects).

Εξηγήσεις

Οι παρατηρήσεις αυτές είναι "δικαιολογημένες" καθώς από το ζητούμενο D είδαμε πως την ψηλότερη επίδραση στη μέση πρόβλεψη την είχαν τα ca , cp .

Παραδείγματα:

Επίσης, αν βάλουμε x = original sample ordering και πάμε στο $x = 0$ (όπου αυτό είναι το sample index των ζητούμενων A και B) βλέπουμε ομοιότητες με τα ζητούμενα A και B . Συγκεκριμένα, παρατηρούμε ότι οι τιμές των γνωρισμάτων είναι πάρα πολύ κοντά σε αυτές του B ερωτήματος (αν αλλάζουμε τον άξονα y στις τιμές που περιγράφονται στο B - oldpeak, slope, ca , cp , thalach).

Όσο για τα χρώματα (μπλε = μείωση πρόβλεψης και κόκκινο = αύξηση) βλέπουμε ότι είναι τα αντίστοιχα γνωρίσματα έχουν τα ίδια χρώματα με αυτά του A και B .

Περιγραφή 5

Εικόνες που χρησιμοποιήθηκαν:

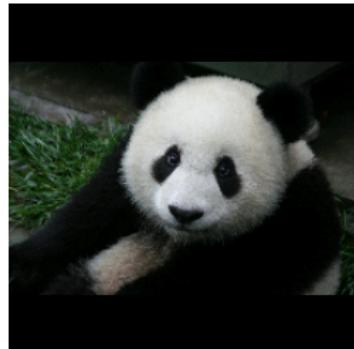
Indian cobra



Axolotl



Giant Panda



Ζητούμενο Α:

```
Model: "sequential_2"
```

Layer (type)	Output Shape	Param #
<hr/>		
inception_v1 (KerasLayer)	(None, 1001)	6633209
<hr/>		
Total params: 6633209 (25.30 MB)		
Trainable params: 0 (0.00 Byte)		
Non-trainable params: 6633209 (25.30 MB)		

Παρατηρούμε πώς στο Inception_v1 το νευρωνικό δίκτυο αναγνωρίζει σωστά και με ακρίβεια 100% το axolotl. Για τα υπόλοιπα όμως, δεν εκφράζει τόσο μεγάλη βεβαιότητα όσον αφορά την αναγνώριση των εικόνων. Η κόμπρα αναγνωρίζεται με ποσοστό επιτυχίας 68.4%, ενώ το panda, το δίκτυο είναι αρκετά βέβαιο με ποσοστό επιτυχίας 89.4%.

```
Model: "sequential_4"
```

Layer (type)	Output Shape	Param #
<hr/>		
inception_v2 (KerasLayer)	(None, 1001)	11199137

```
=====
Total params: 11199137 (42.72 MB)
Trainable params: 0 (0.00 Byte)
Non-trainable params: 11199137 (42.72 MB)
```

Όσο αφορά το Inception_v2 το νευρωνικό δίκτυο εκφράζει αρκετά μεγαλύτερη βεβαιότητα όσον αφορά την αναγνώριση των εικόνων σε σχέση με το inception_v1. Η κόμπρα αναγνωρίζεται με ποσοστό επιτυχίας 91.7%, που είναι πολύ καλύτερο από το inception_v1, ενώ όσον αφορά το panda, το δίκτυο είναι και πάλι αρκετά βέβαιο με ποσοστό επιτυχίας 93.6%. Αντίθετα όμως για το axolotl, δεν αναγνωρίζεται όπως το inception_v1 (100%) και η πρόβλεψη έπεισε στο 94.2%.

```
Model: "sequential_5"

Layer (type)          Output Shape         Param #
=====
inception_v3 (KerasLayer)    (None, 1001)      23853833
=====

Total params: 23853833 (91.00 MB)
Trainable params: 0 (0.00 Byte)
Non-trainable params: 23853833 (91.00 MB)
```

Στο Inception_v3 παρατηρούμε μια σημαντική αύξηση της βεβαιότητας σε σχέση με τα υπόλοιπα 2 versions. Αρχικά αναγνωρίζει σχεδόν άριστα την κόμπρα με 99.6% (πολύ καλύτερο σε σχέση με τα προηγούμενα) ενώ το axolotl και το panda με 100% επιτυχία.

```
Model: "sequential_6"

Layer (type)          Output Shape         Param #
=====
inception_resnet_v2 (Keras    (None, 1001)      55875273
Layer)
=====

Total params: 55875273 (213.15 MB)
Trainable params: 0 (0.00 Byte)
```

```
Non-trainable params: 55875273 (213.15 MB)
```

Στην έκδοση Inception_resnet_v2, παρατηρούμε τα υψηλότερα ποσοστά επιτυχίας μέχρι στιγμής, καθώς όλα τα ζώα αναγνωρίζονται σωστά με ποσοστό "σιγουρίας" 100%.

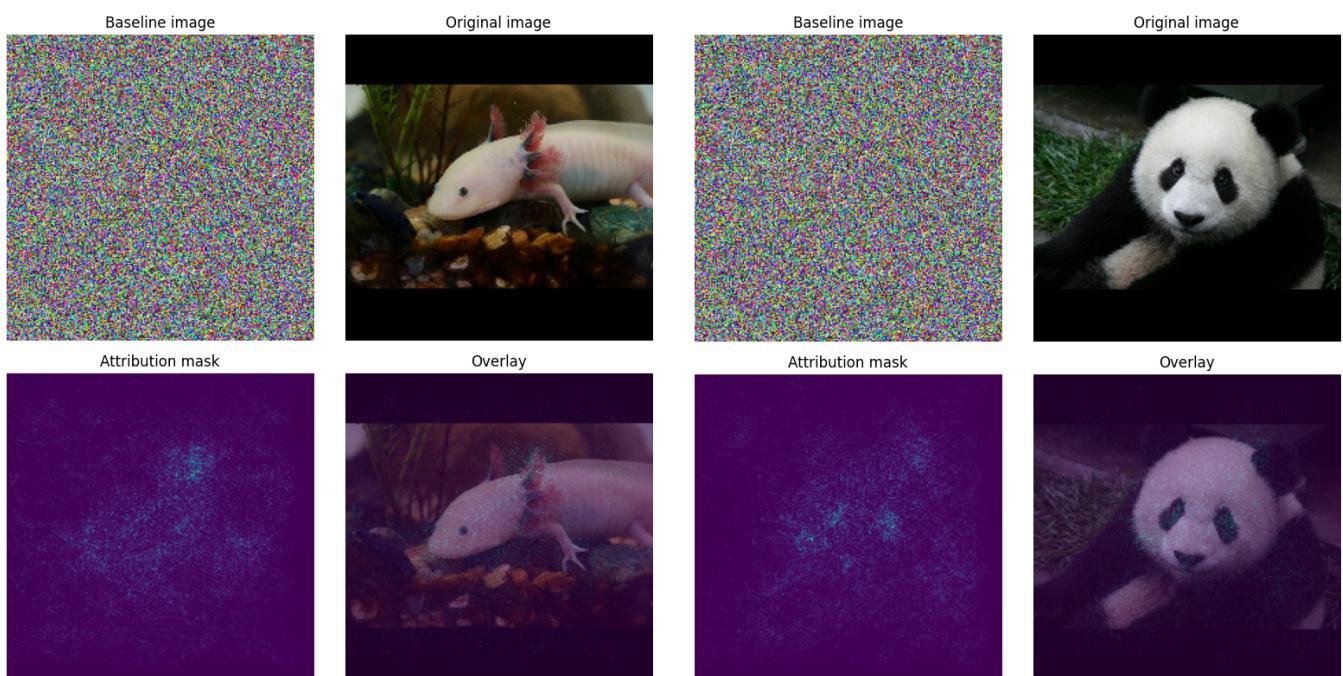
Ζητούμενο Β:

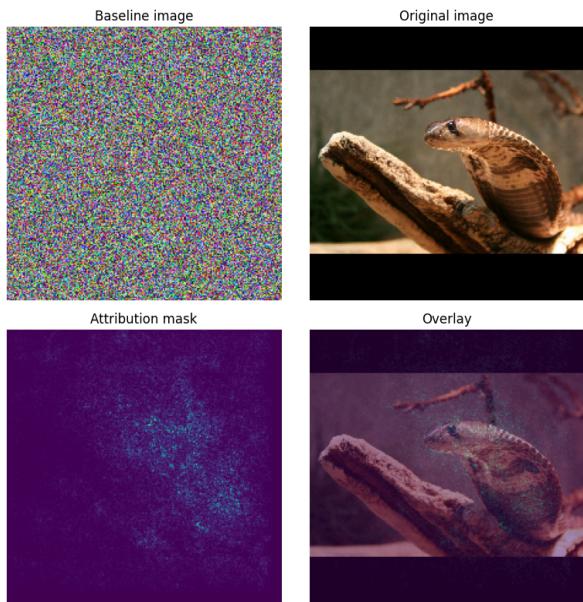
Παρατηρούμε ότι το καλύτερο αποτέλεσμα (classification) για κάθε εικόνα το έχει το inception_resnet_v2, οπότε και το επιλέγουμε. Ακολουθεί η δοκιμή με baseline random:

```
Model: "sequential_9"
```

Layer (type)	Output Shape	Param #
<hr/>		
inception_resnet_v2 (Keras Layer)	(None, 1001)	55875273
<hr/>		
Total params:	55875273 (213.15 MB)	
Trainable params:	0 (0.00 Byte)	
Non-trainable params:	55875273 (213.15 MB)	

Αποτελέσματα:





Χρησιμοποιώντας το τυχαίο baseline, παρατηρούμε ότι το μοντέλο δίνει έμφαση σε διάφορα pixels (που επισημάνονται με ανοιχτό χρώμα στο attribution mask) με αποτέλεσμα οι παραγόμενες εξήγησεις (attribution mask, overlay) να έχουν πιο πολύ "θόρυβο" απ'ότι προηγουμένως που χρησιμοποιήσαμε το "μαύρο" baseline.

Μόνο στο attribution mask για το giant panda βλέπουμε ότι το μοντέλο έχει δώσει έμφαση σε pixels που αναδεικνύουν τα χαρακτηριστικά του ζώου (όπως το πρόσωπο και τα μάτια).

Στη εικόνα του axolotl από την άλλη, βλέποντας το overlay παρατηρούμε ότι το μοντέλο έχει δώσει σημασία με γενικό τρόπο σε pixels πάνω στο σώμα του ζώου, χωρίς όμως να τονίζει κάποια ιδιαίτερα χαρακτηριστικά του (όπως όταν χρησιμοποιούσαμε το μαύρο baseline στο inception_resnet_v2 το οποίο έδινε σημασία σε χαρακτηριστικά όπως τα πτερύγια - πόδια του axolotl). Όμοια συμπεριφορά μπορούμε να παρατηρήσουμε και για την εικόνα του indian cobra.

Μπορείτε να δείτε όλο το κώδικα του colab [εδώ](#).