

ΣΥΣΤΗΜΑΤΑ ΥΠΟΣΤΗΡΙΞΗΣ ΑΠΟΦΑΣΕΩΝ

Χρήστος Μπούμπουρας
Στέλιος Γαβράς

ΠΕΡΙΓΡΑΦΗ

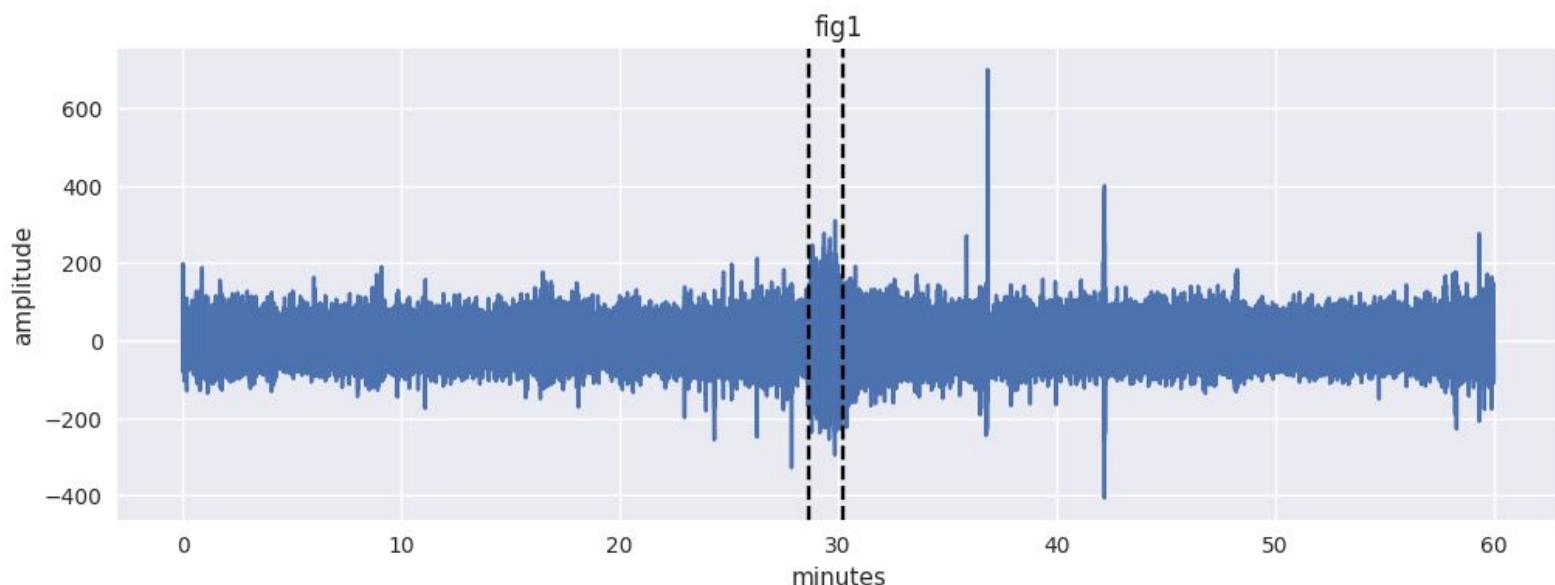
Χρησιμοποιήσαμε τεχνικές supervised machine learning και πιο συγκεκριμένα regression για να προβλέψουμε επιληπτικά γεγονότα.

Η βάση από την οποία πήραμε τα δεδομένα μας περιέχει καταγραφές EEG για 22 άτομα, στα οποία έχουν γίνει από 9-42 πειράματα με έναν μικρό αριθμό πειραμάτων που περιέχουν ένα ή παραπάνω επεισόδια επιληψίας. Καθε πείραμα περιέχει περίπου 1 ωρα δεδομένων από 22 διαφορετικά κανάλια με συχνότητα 256 τιμών ανα δευτερόλεπτο.

Κάναμε δυο τεστ, στο πρώτο χρησιμοποιήσαμε την μέθοδο μας στα πειράματα με επιληψία του πρώτου ασθενή και το δεύτερο σε τυχαία πειράματα με επιληψία από τους πρώτους 10 ασθενείς.

ΜΕΘΟΔΟΣ

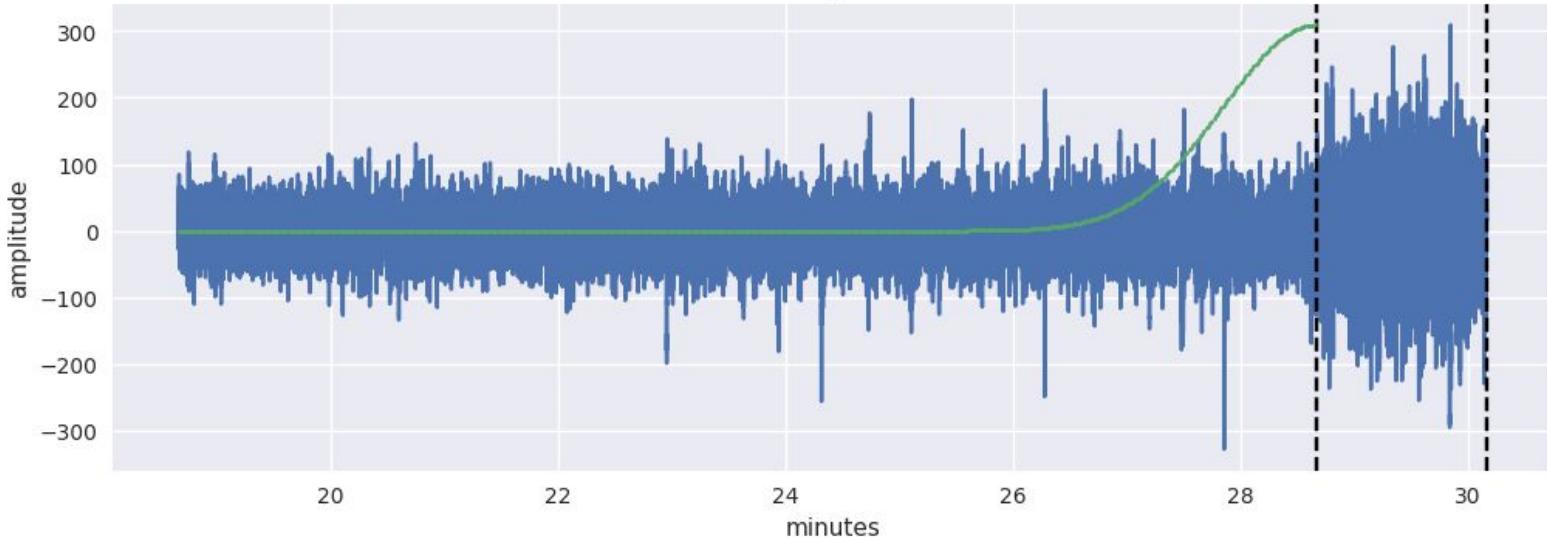
Στο fig1 βλέπουμε ένα ενδεικτικό πείραμα από τον πρώτο ασθενή στο κανάλι 12 και την περίοδο ύπαρξης επιληψίας.



Στη συνέχεια όπως φαίνεται στο fig2 επιλέγουμε ενα διάστημα 10 λεπτών πριν την επιληψία και δίνουμε μια πιθανότητα ύπαρξης επιληψίας(πράσινη γραμμή). Αυτή η

πιθανότητα επιλέξαμε να ανεβαίνει μετά τα 2 λεπτά πριν την έναρξή της, με πιθανότητα ίση με 1 στο σημείο που ξεκινάει η επιληψία.

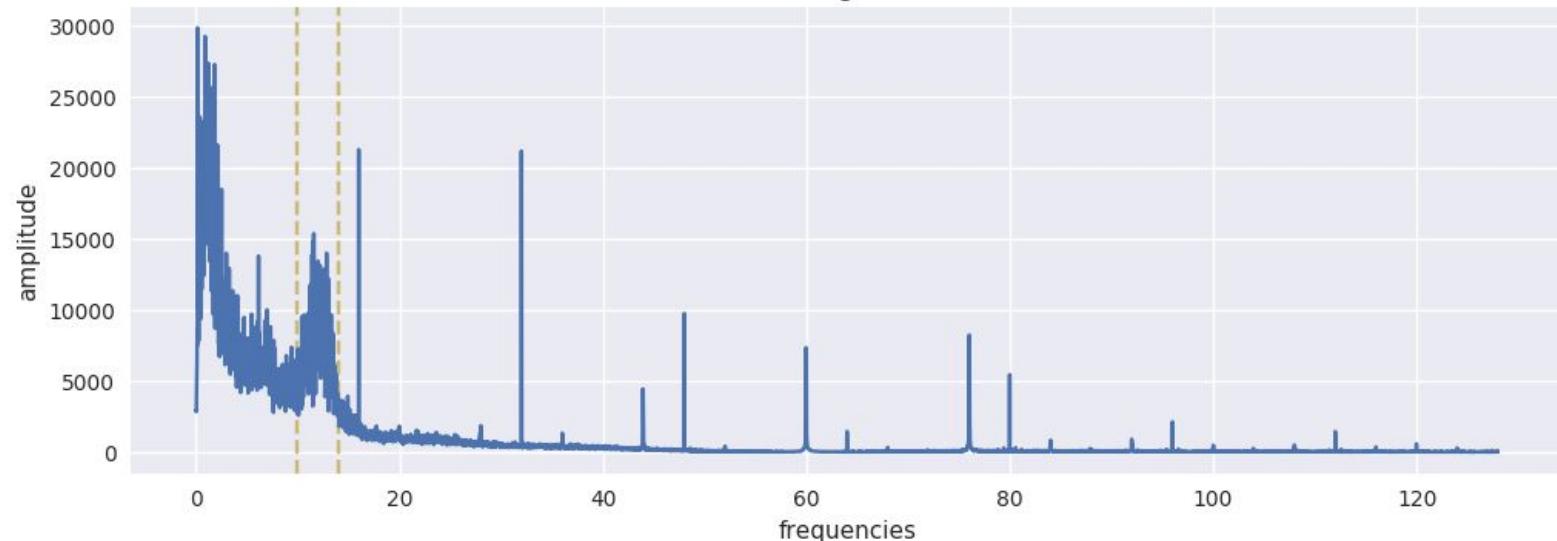
fig2



Μέσα σε αυτά τα 10 λεπτά επιλέγουμε διαστήματα του ενός λεπτού ανα δευτερόλεπτο μεχρι την έναρξη της επιληψίας και σε κάθε διάστημα χρησιμοποιούμε τις παρον συχνότητες για να βγάλουμε κάποια συμπεράσματα.

Πιο συγκεκριμένα αυτό γίνεται σε κάθε πείραμα και για κάθε κανάλι, για παράδειγμα

fig3



επιλέγοντας μια συγκεκριμένη συχνότητα από 10-14(fig3) για καθε μονόλεπτο παράθυρο που παίρνουμε καταλήγουμε με τον παρακάτω ενδεικτικό πίνακα. Εχουμε για στήλες τα διαφορετικά κανάλια και την πιθανότητα ύπαρξης επιληψίας(y) και για γραμμές τα

παραθυρά ενός λεπτού που παίρνουμε ανα δευτερόλεπτο, με τις αντίστοιχες τιμές που προκύπτουν αμα βγάλουμε τον μέσο όρο του amplitude των συχνοτήτων από 10-14.

	0	1	2	3	4	5	6	7	8	9	...	14	15	16	17	18	19	20	21	22	y
0	1683	1694	1973	2185	2291	2172	1902	2373	1943	1924	...	2031	2464	2622	2956	1973	1295	2346	1218	2031	2.200899e-64
1	1679	1692	1978	2179	2287	2169	1895	2377	1938	1938	...	2041	2464	2647	2941	1978	1272	2363	1218	2041	2.717504e-61
2	1674	1698	1972	2159	2304	2178	1871	2357	1946	1918	...	2020	2426	2624	2926	1972	1283	2387	1206	2020	9.923116e-60
3	1695	1715	1966	2185	2314	2205	1866	2377	1953	1923	...	2003	2437	2610	2936	1966	1296	2320	1199	2003	1.615114e-58
4	1680	1700	1984	2118	2320	2210	1890	2304	1953	1916	...	2013	2405	2597	2964	1984	1280	2274	1191	2013	1.714784e-57
5	1669	1664	1953	2101	2315	2212	1860	2305	1943	1892	...	1995	2370	2614	2937	1953	1267	2263	1198	1995	1.381125e-56
6	1668	1631	1937	2115	2315	2191	1877	2294	1954	1896	...	2014	2415	2605	2955	1937	1248	2251	1201	2014	9.127075e-56
7	1657	1620	1905	2123	2310	2186	1835	2316	1956	1891	...	2000	2445	2603	2898	1905	1237	2265	1189	2000	5.186953e-55
8	1637	1599	1955	2148	2256	2230	1848	2329	1962	1907	...	2009	2434	2634	2936	1955	1225	2242	1172	2009	2.614307e-54
9	1636	1587	1939	2160	2132	2155	1841	2341	1857	1877	...	2030	2428	2553	2878	1939	1226	2263	1165	2030	1.194005e-53

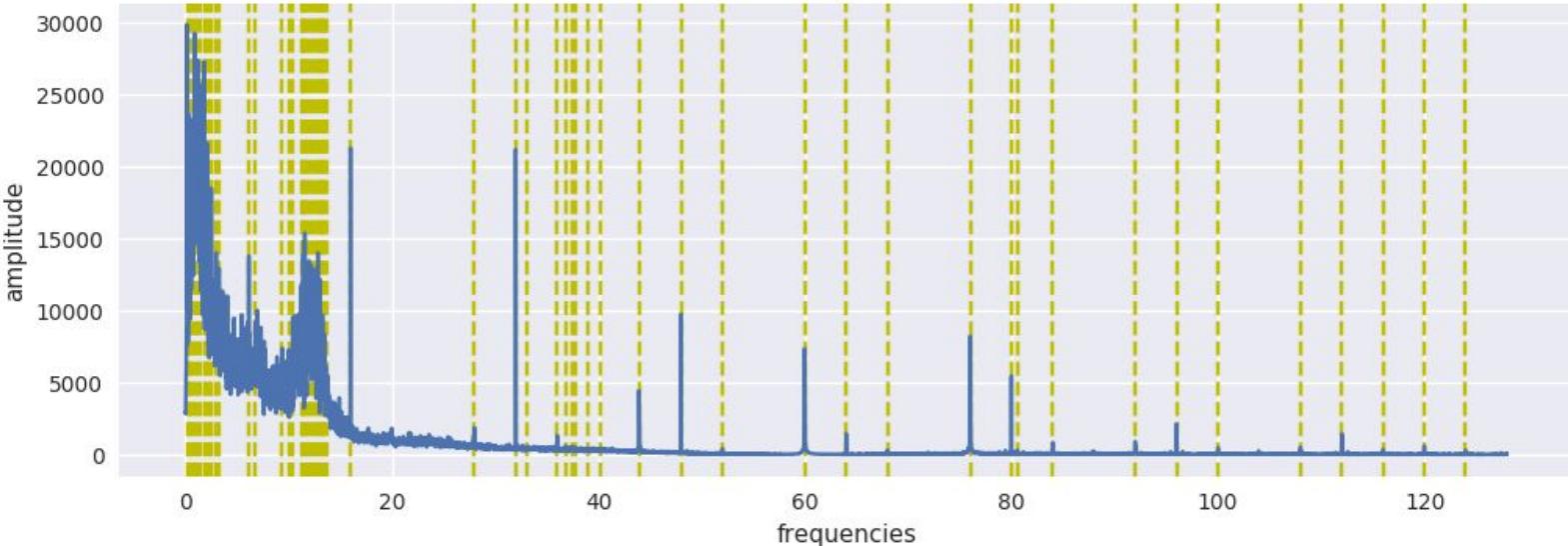
ΑΠΟΤΕΛΕΣΜΑΤΑ

Στα αποτελέσματά μας δοκιμάσαμε ένα μεγάλο εύρος συχνοτήτων και επιλεξαμε τις πιο αποδοτικές. Η επιλογή όλων αυτών των συχνοτήτων έγινε με δύο τρόπους, ο ένας ήταν παράγοντας πολλά τυχαία διαστήματα και ο άλλος επιλέγοντας τις πιο ισχυρές.(fig4,fig5)

(fig4)

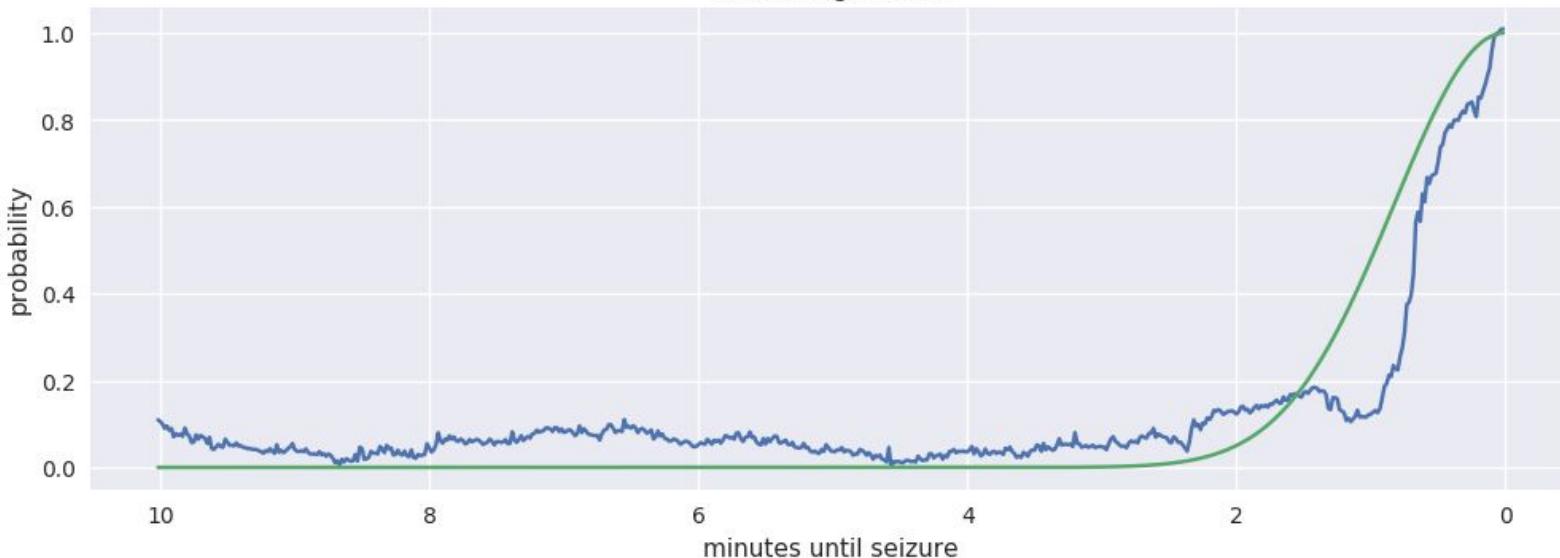
[2-104], [17-41], [41-89], [22-52], [52-65], [65-105], [5-39], [39-51], [51-88], [88-116], [6-26], [26-47], [47-74], [74-94], [94-117], [12-22], [22-31], [31-52], [52-63], [63-76], [76-98], [98-107], [0-23], [23-35], [35-37], [37-48], [48-68], [68-76], [76-95], [95-100], [100-108], [2-6], [6-16], [16-22], [22-29], [29-30], [30-38], [38-43], [43-49], [49-58], [58-61], [61-67], [67-74], [74-78], [78-88], [88-91], [91-96], [96-107], [107-109], [109-115], [0-6], [6-9], [9-12], [12-16], [16-20], [20-26], [26-29], [29-34], [34-38], [38-40], [40-46], [46-49], [49-52], [52-56], [56-60], [60-66], [66-71], [71-74], [74-79], [79-82], [82-85], [85-90], [90-93], [93-97], [97-101], [101-105], [105-109], [109-115], [115-116],

fig5

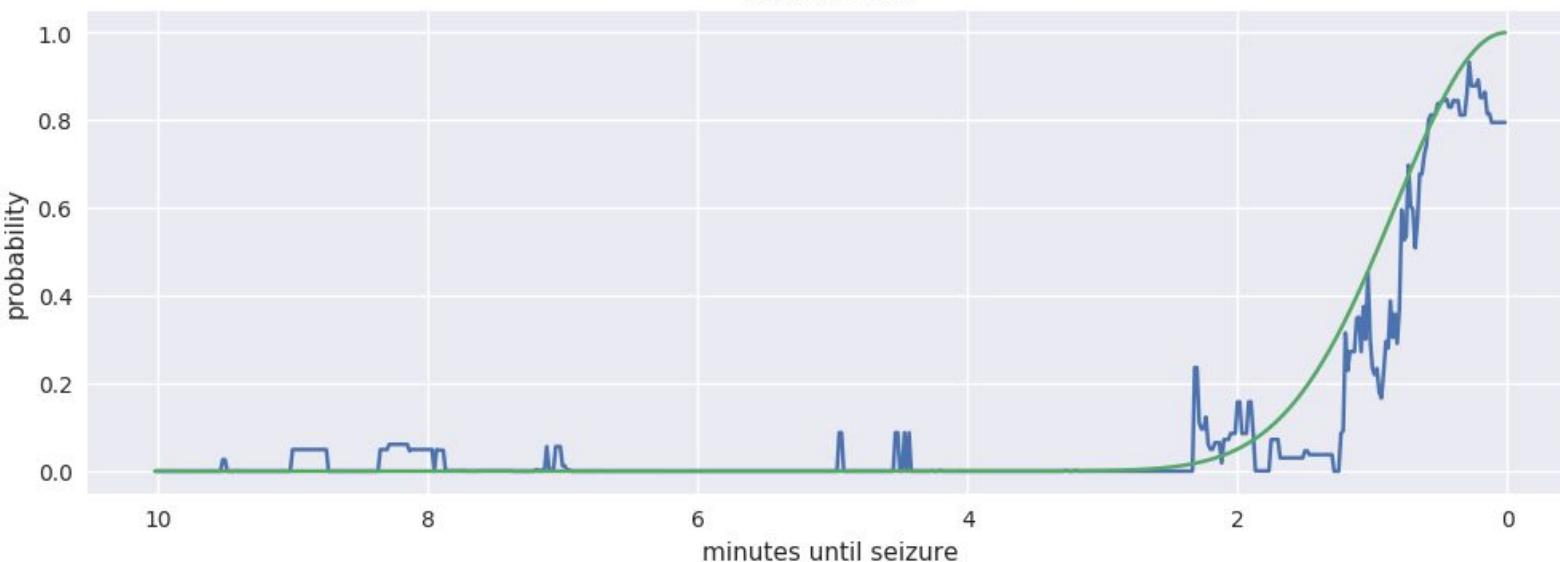


Στο πρώτο τεστ χρησιμοποιήσαμε τα 6 πειράματα με επιληπτικά γεγονότα του πρώτου ασθενή, όπου κάναμε train στα πέντε και τεστάραμε σε ένα έκτο. Για την αξιολόγηση των αποτελεσμάτων συνδυάσαμε τα R squared , το Mean Absolute Error και το Mean Squared Error και επιλέξαμε τις συχνότητες με τις καλύτερες μετρήσεις. Στα παρακάτω figures φαίνονται ενδεικτικά οι μέσες τιμές των καλύτερων αποτελεσμάτων διότι ήταν πάρα πολλά. Με πράσινη γραμμή είναι οι πραγματικές τιμές και με μπλέ οι προβλεπόμενες.

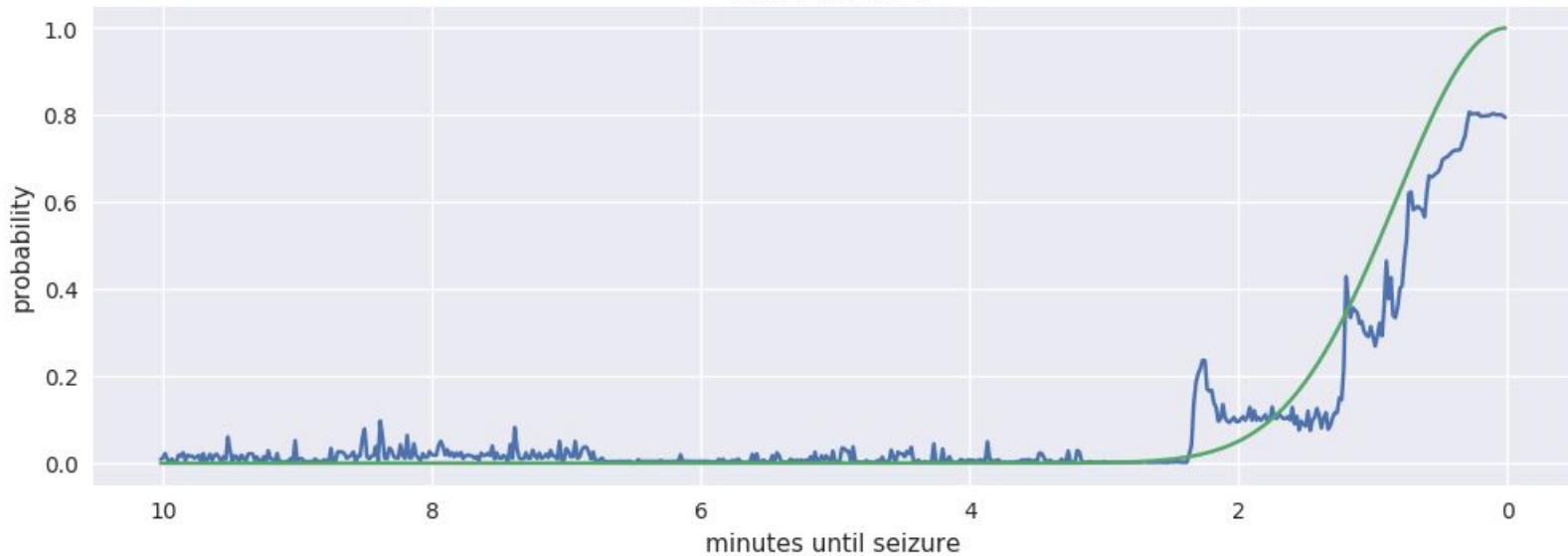
LinearRegression



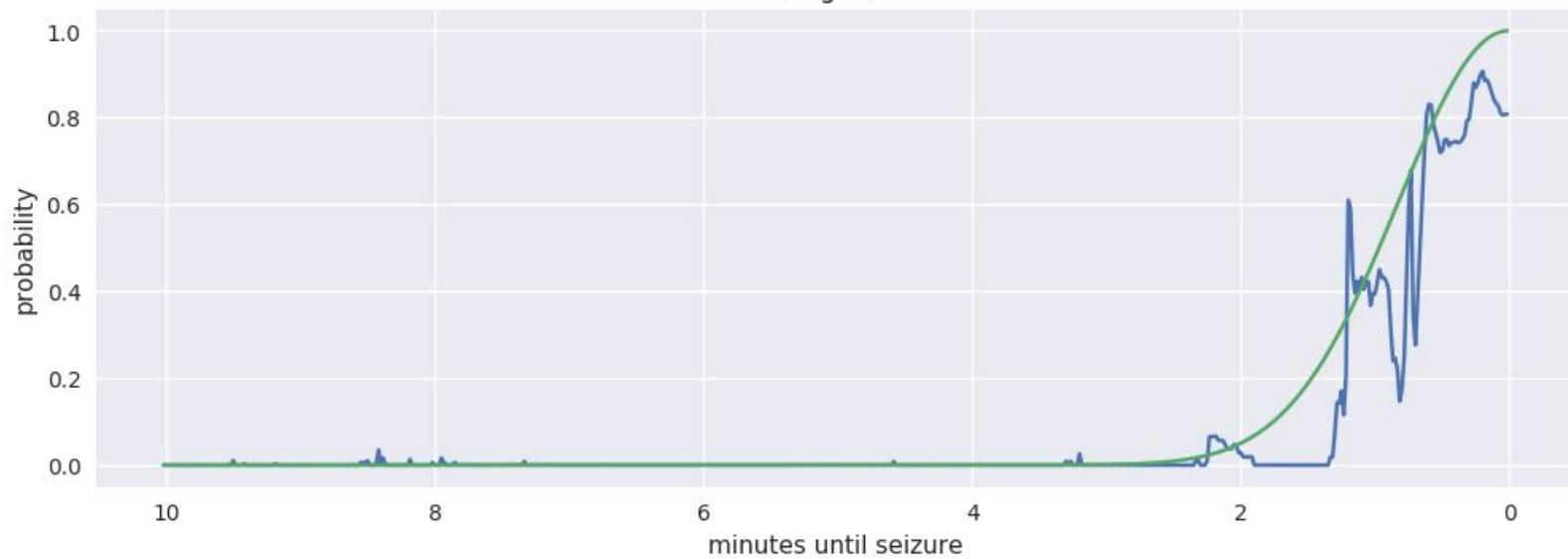
DecisionTree



RandomForest



KNeighbors



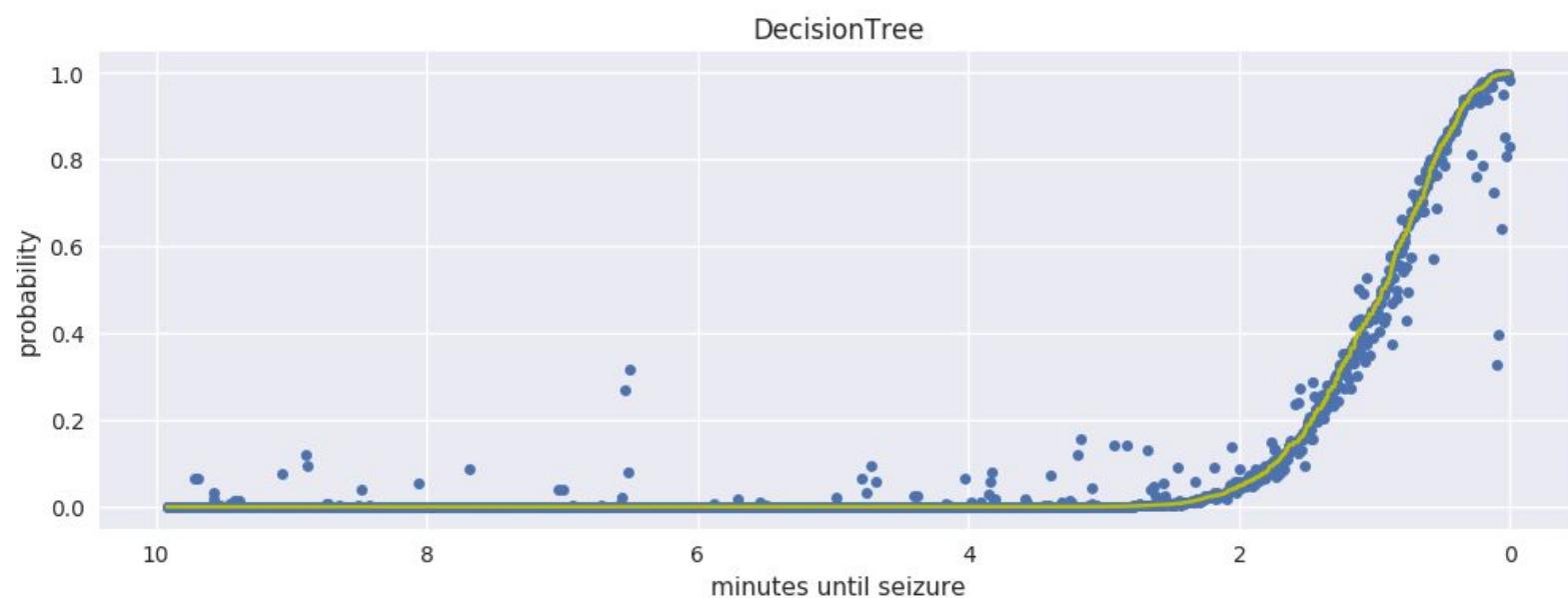
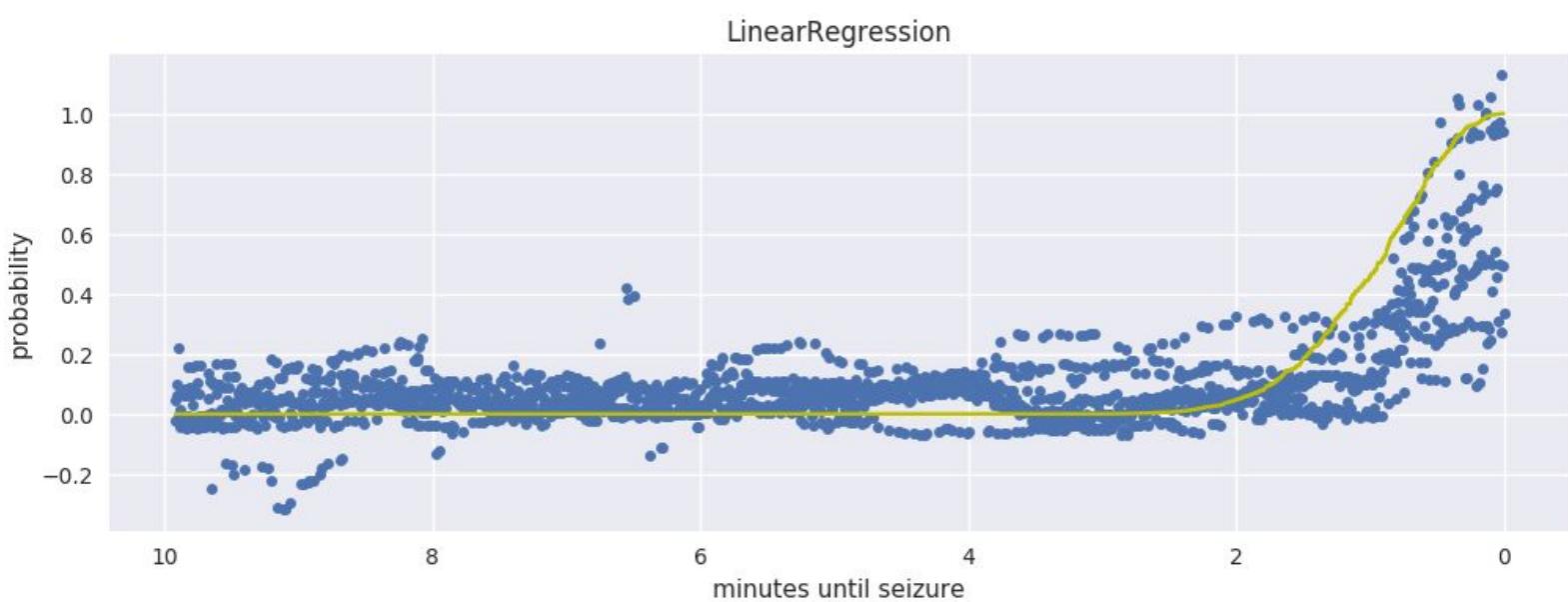
Στο παρακάτω πινακάκι φαίνονται οι αντίστοιχες μέσες τιμές των μέτρων αξιολόγησης.

	LinearRegression	DecisionTree	RandomForest	KNeighbors
R	0.916322	0.886033	0.949655	0.947432
MAE	0.067547	0.046392	0.039046	0.032483
MSE	0.011957	0.014769	0.007808	0.007353

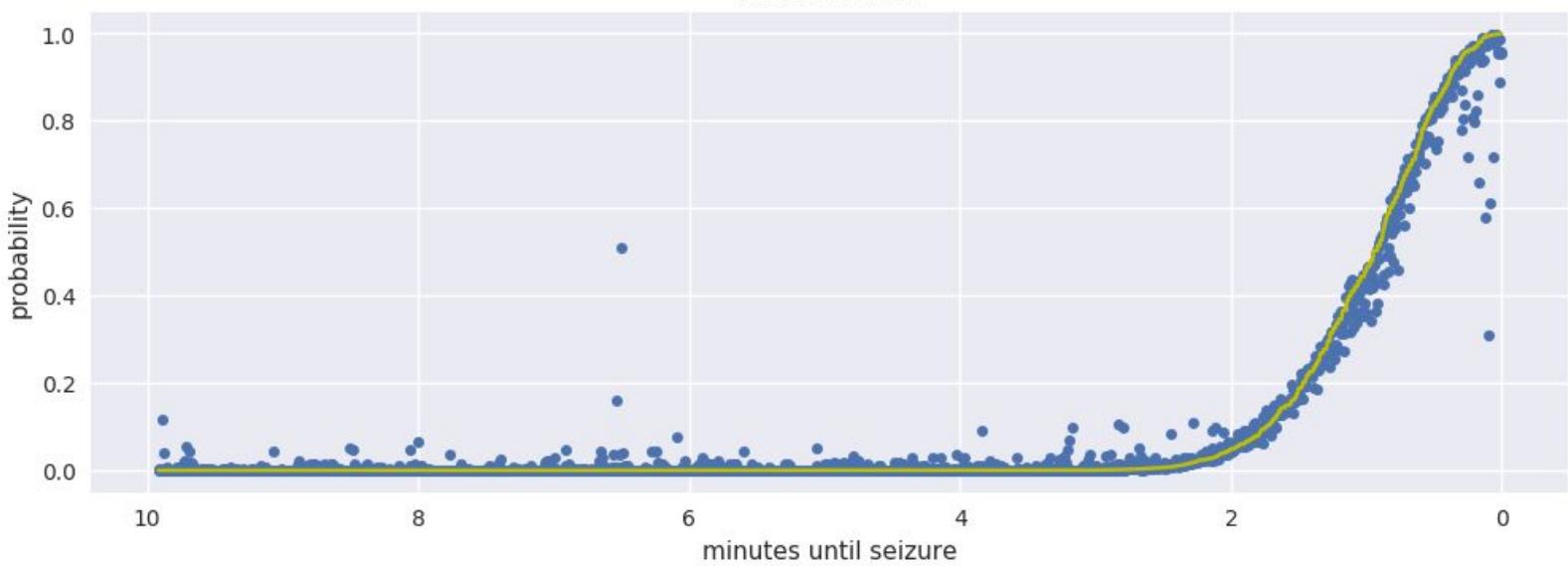
Για το δεύτερο τέστ χρησιμοποιήσαμε τυχαία πειράματα από δέκα διαφορετικούς ασθενείς. Σε αντίθεση με την προηγούμενη περίπτωση, τα δεδομένα ήταν τυχαία κατανεμημένα και η εκπαίδευση έγινε στο 70% των δεδομένων καθώς και η πρόβλεψη στο υπόλοιπο 30%.

Οπως φαίνεται κι εδώ τα αποτελέσματα είναι πάρα πολύ καλα πέρα από το Linear regression.

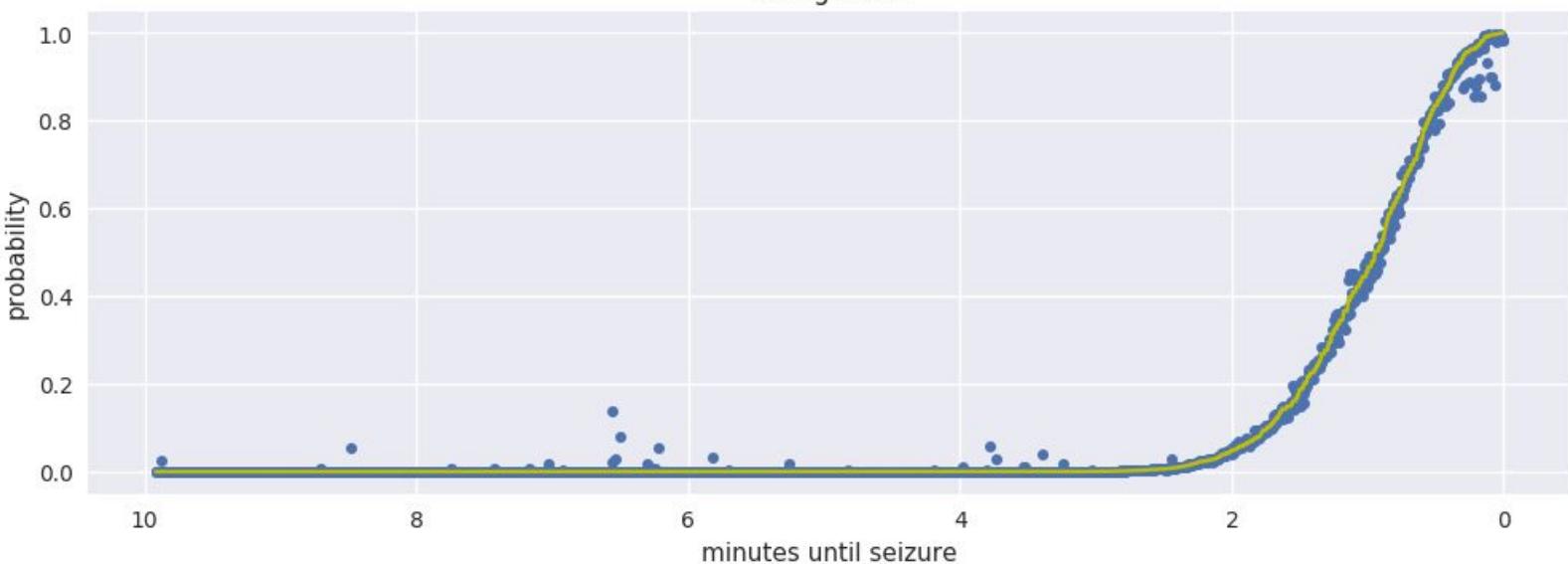
	LinearRegression	DecisionTree	RandomForest	KNeighbors
R	0.711274	0.974313	0.984801	0.997374
MAE	0.114993	0.008962	0.010387	0.004232
MSE	0.030692	0.003178	0.001913	0.000329



RandomForest



KNeighbors



Η συγκεκριμένη ανάλυση έχει ακόμα μεγάλο περιθώριο βελτίωσης και εξέλιξης σε κάπποιου είδους DSS διότι είναι φανερό ότι υπάρχει μεγάλο ποσοστό επιτυχίας.