

Ανάπτυξη Λογισμικού για Αλγοριθμικά Προβλήματα - Εργασία 3

Κωνσταντína Έλληνα - 1115201600046

Απόστολος Λύρας - 1115201600097

Παρατηρήσεις για Ερώτημα Β

Τα πειράματα που κάναμε για το συγκεκριμένο ερώτημα αφορούν τις υπερπαραμέτρους που ζητούνται από την εκφώνηση και παρακάτω παραθέτουμε τις παρατηρήσεις μας. Τα πειράματα έχουν γίνει για $n = 1$ και έχουμε βάλει early stopping, έτσι ώστε να σταματάει το μοντέλο όταν δεν έχει βελτιωθεί το validation loss για 5 συνεχόμενες φορές. Το αρχείο των χρονοσειρών που χρησιμοποιήσαμε είναι το 'nasdaq2007_17.csv'. Ως default τιμή για το error value βάλαμε το 0.2 που είδαμε ότι βρίσκει διάφορες ανωμαλίες.

Το ερώτημα αυτό τρέχει για tensorflow version 2.1. Είχαμε κάποιο πρόβλημα με το συγκεκριμένο και έπρεπε να αλλάξουμε κάποια πράγματα, όπως το `tf.executing_eagerly()`.

Delay

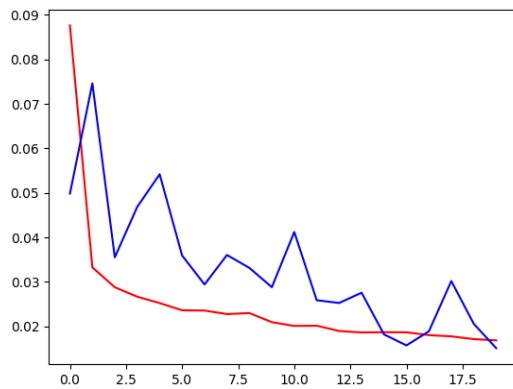
Αρχικά θέλαμε να δούμε τις διαφορές που θα έχει το εκτελέσιμο με διαφορετικό delay. Γι' αυτό το λόγο αφήσαμε σταθερές τις υπόλοιπες παραμέτρους και είδαμε τα εξής:

Με `lstm_units = 64`, `lstm_layers = 1`, `dropout = 0.1`, `batch_size = 32`, `max_epochs = 20`

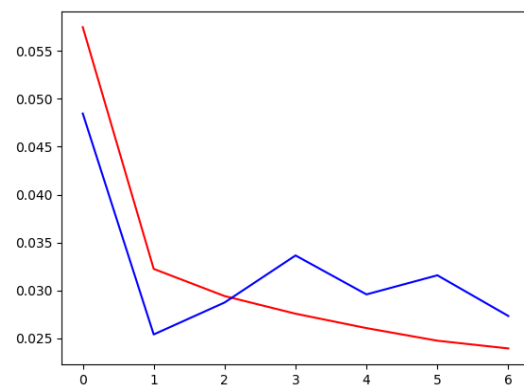
Delay	10	50	100	150	200	300
Test loss	0.03525	0.05833	0.10244	0.09907	0.11509	0.14633

Το **validation loss** βελτιώνεται κάπως με μεγαλύτερο delay όπως φαίνεται παρακάτω, όπου Blue line = validation loss, Red line = Train loss.

Delay = 1

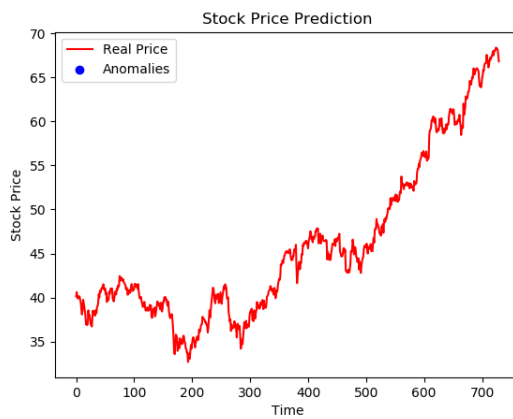


Delay = 200

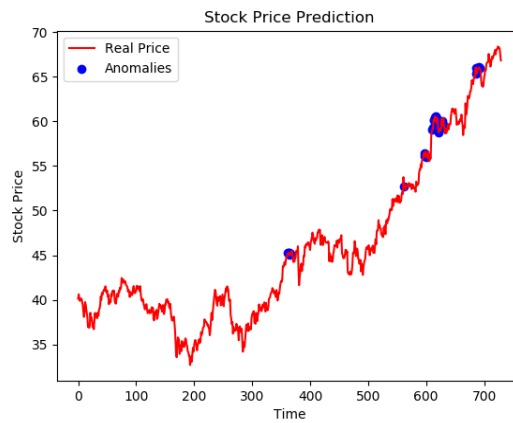


Παρακάτω φαίνεται ότι για μικρό delay δεν βρίσκει ανωμαλίες. Για 10 και για 50 που δοκιμάσαμε εμείς δεν βρίσκει κάποιο σημείο. Όσο μεγαλώνει τόσο περισσότερες ανωμαλίες βρίσκει το μοντέλο.

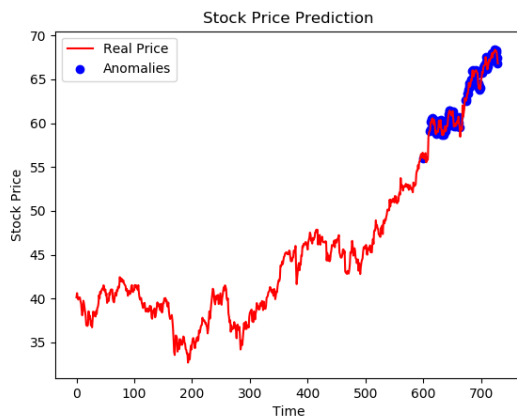
Delay = 10



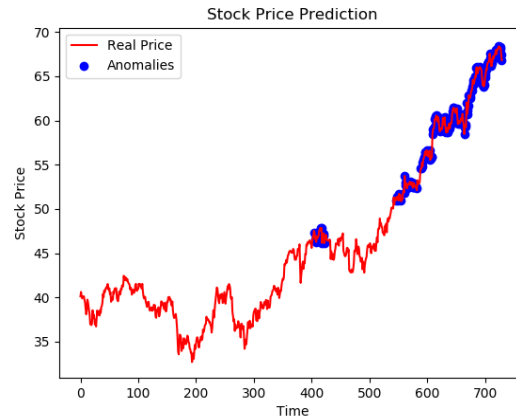
Delay = 100



Delay = 200



Delay = 300



Διαλέγουμε ως βέλτιστη τιμή το 200 έτσι ώστε να βρίσκει κάποιες ανωμαλίες.

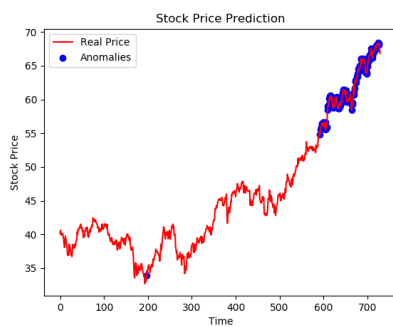
LSTM units

Με lstm_layers = 1, dropout = 0.1, max_epochs = 20, batch_size = 32, delay = 200

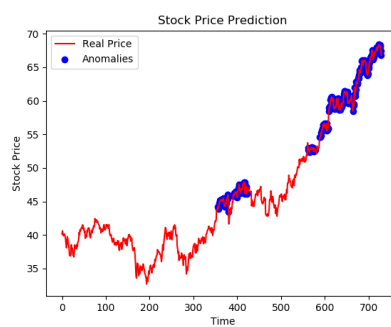
Units	8	16	32	64	128
Test loss	0.09788	0.12473	0.12501	0.14631	0.11530

Βλέπουμε ότι για μικρό αριθμό units δεν βρίσκει ανωμαλίες. Θεωρούμε ότι το 64 είναι το καλύτερο.

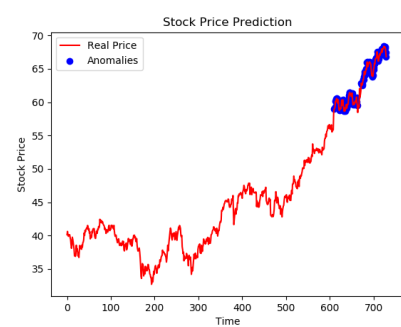
Units = 16



Units = 64



Units = 128



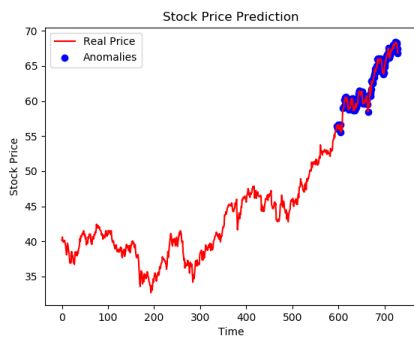
LSTM layers

Με lstm_units = 64, dropout = 0.1, max_epochs = 20, batch_size = 32, delay = 200

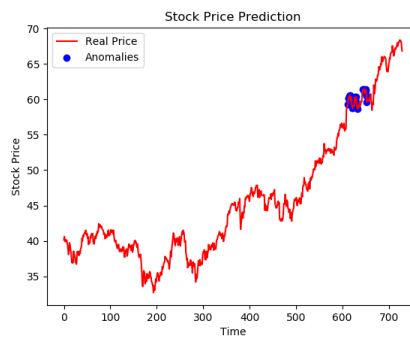
Layers	1	2	3
Test loss	0.11822	0.10531	0.10268

Για διαφορετικά layers έχουμε και διαφορετικές ανωμαλίες:

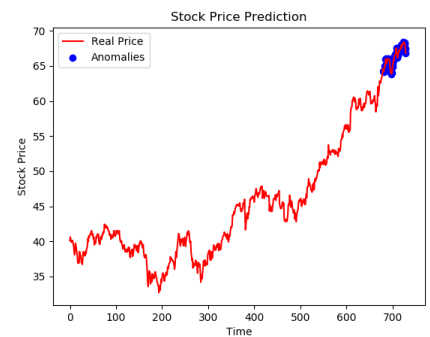
Layers = 1



Layers = 2



Layers = 3



Epochs

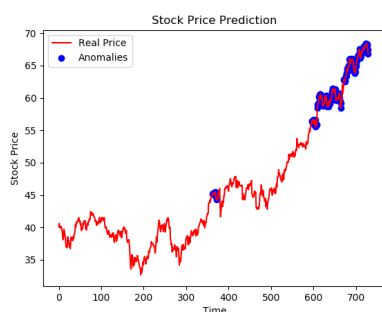
Με lstm_units = 64, lstm_layers = 1, dropout = 0.1, batch_size = 32, delay = 200

Όσο περισσότερες εποχές τόσο πιο πολύ μειώνεται το loss, καθώς κάνει περισσότερα περάσματα στο train set.

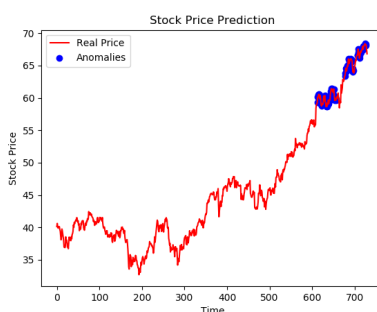
Epochs	5	10	20	40	50
Test loss	0.12263	0.11105	0.12104	0.11996	0.14247

Τα αποτελέσματα μοιάζουν μεταξύ τους. Παρατηρήσαμε ότι στις 40 εποχές έκανε early stopping στην 25η, ενώ στις 50 εποχές σταμάτησε στις 37. Από εκεί καταλαβαίνουμε ότι δεν χρειάζονται τόσες εποχές τελικά, οπότε στο μοντέλο μας χρησιμοποιούμε 20 εποχές.

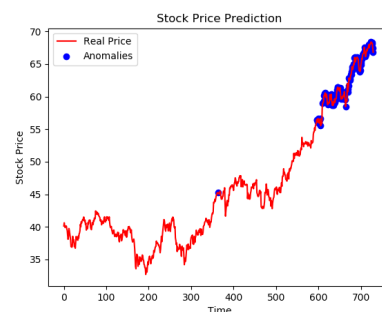
Epochs = 5



Epochs = 10



Epochs = 40

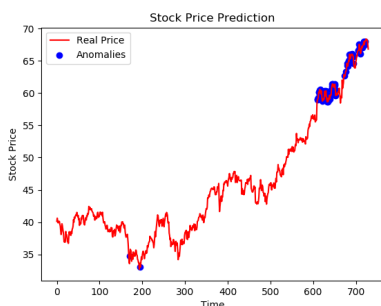


Batch Size

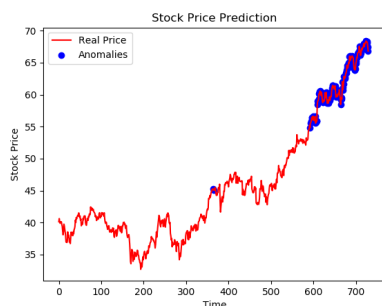
Με lstm_units = 64, lstm_layers = 1, dropout = 0.1, epochs = 20, delay = 200

Batch size	8	32	64	128
Test loss	0.10877	0.12809	0.15218	0.11474

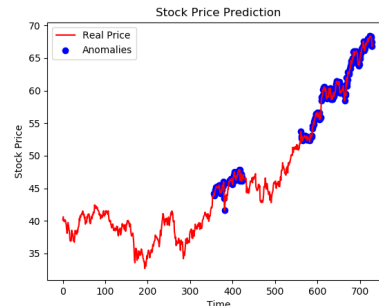
Batch_size = 8



Batch_size = 32



Batch_size = 64



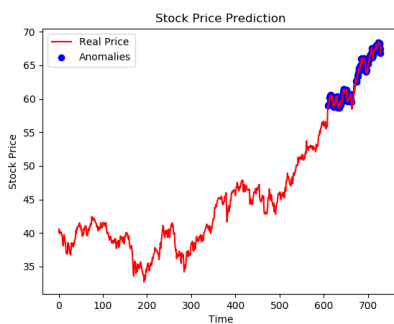
Dropout

Με $\text{lstm_units} = 64$, $\text{lstm_layers} = 1$, $\text{epochs} = 20$, $\text{batch size} = 32$, $\text{delay} = 200$

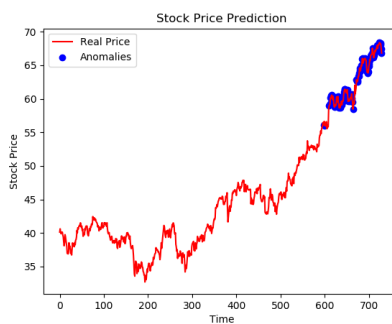
Dropout	0.05	0.1	0.5	0.9
Test loss	0.11274	0.11593	0.09866	0.23319

Βλέπουμε τη διαφορά του dropout όσο αυξάνεται. Η τιμή του dropout είναι $[0,1)$.

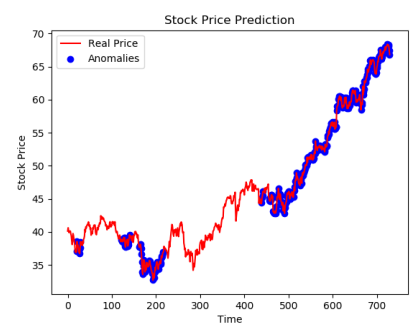
Dropout = 0.05



Dropout = 0.1



Dropout = 0.9



Optimal Model

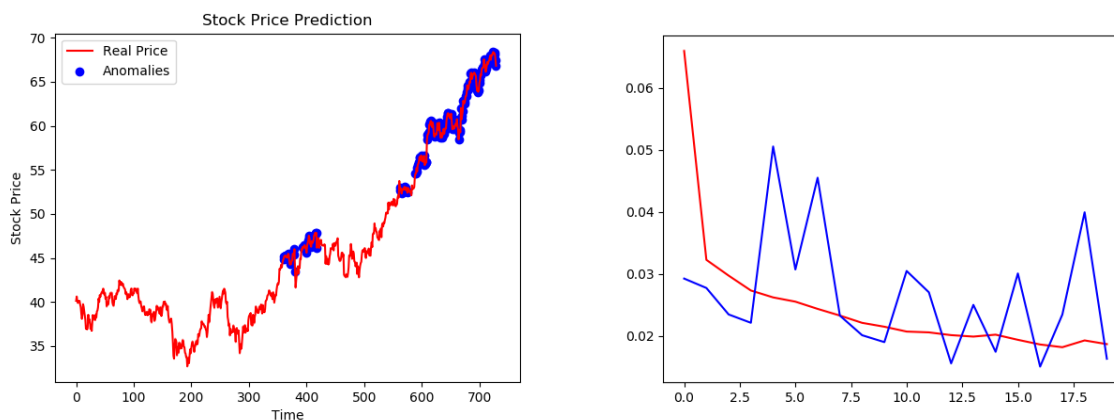
Ως βέλτιστο μοντέλο έχουμε βάλει τις παρακάτω υπερπαραμέτρους:

- $\text{Lstm_units} = 64$
- $\text{Lstm_layers} = 1$
- $\text{Dropout} = 0.1$
- $\text{Epochs} = 20$
- $\text{Batch_size} = 32$
- $\text{Delay} = 200$

Το αποτέλεσμα που πήραμε για το **loss** είναι:

Test sample 0 average loss: 0.14234763007995957

Το αποτέλεσμα του plot για τα **anomalies** είναι στα αριστερά και το αποτέλεσμα του **validation loss** είναι στα δεξιά:



Optimal Model for $n>1$

Τρέξαμε το ερώτημα και με $n = 3$ και πήραμε τα παρακάτω αποτελέσματα:

Test sample 0 average loss: 0.04758063093659463

Test sample 1 average loss: 0.06119170404044848

Test sample 2 average loss: 0.07405089093277102

Για $n>1$ δεν βρίσκει anomalies όπως είδαμε στα διαγράμματα.

Στο φάκελο του ερωτήματος Β υπάρχουν στους αντίστοιχους φακέλους όλα τα αποτελέσματα που βγάλαμε, τα οποία δείχνουμε παραπάνω.