|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| שם הסקריפט | L1 regulation | L2 regulation | Drop  out | #History points | #Epochs | Learning  Rate | Batch size | LSTM  L1 | LSTM  L2 | DENSE  L1 | Accuracy 1-MAPE)) | זמן ריצה\*\* (דקות) |
| Model\_baseline | 0 | 0 | 0.2 | 10 | 50 | 0.00005 | 512 | 50 | 0 | 64 | 70% | 25 |
| model\_L1\_Regulation\_Dropout | 0.0002 | 0 | 0.2 | 10 | 50 | 0.00005 | 512 | 50 | 0 | 64 | 58% | 29 |
| model\_L2\_Regulation\_Dropout | 0 | 0.0002 | 0.2 | 10 | 50 | 0.00005 | 512 | 50 | 0 | 64 | 79% | 40 |
| model\_L1\_N\_L2 | 0.0002 | 0.0002 | 0.2 | 10 | 100 | 0.00005 | 512 | 30 | 0 | 64 | 80% | 15 |
| model\_L1\_Regulation\_Nodropout | 0.0002 | 0 | 0 | 10 | 50 | 0.00005 | 512 | 50 | 0 | 64 | 98% | 50 |
| model\_L2\_Regulation\_Nodropout \_moreEpoc | 0 | 0.0002 | 0 | 10 | 100 | 0.00005 | 512 | 50 | 0 | 64 | 94% | 100 |
| model\_L2\_Regulation\_Nodropout \_moreEpoc\_learnrate | 0 | 0.0002 | 0 | 10 | 100 | 0.00002 | 512 | 50 | 0 | 64 | 89% | 50 |
| model\_L2\_Regulation\_Nodropout \_126\_batchsize\_0.002\_L2 | 0 | 0.002 | 0 | 10 | 50 | 0.00005 | 126 | 50 | 0 | 64 | 92% | 30 |
| model\_L1\_Regulation\_Nodropout \_128\_batchsize | 0.0002 | 0 | 0 | 10 | 100 | 0.00002 | 128 | 50 | 0 | 64 |  |  |
| model\_L2\_Regulation\_lstm\_layer1 \_150 | 0 | 0.0002 | 0 | 10 | 100 | 0.00005 | 512 | 150 | 0 | 64 | 93% | 40 |
| model\_L1\_Regulation\_Nodropout \_lstm\_layer\_150 | 0.0002 | 0 | 0 | 10 | 50 | 0.00002 | 512 | 150 | 0 | 64 | 88% | 40 |
| model\_L1\_Regulation\_Nodropout \_lstm\_layer\_150 | 0.0002 | 0 | 0 | 10 | 100 | 0.00005 | 512 | 150 | 0 | 64 | 86% | 30 |
| model\_L2\_Regulation\_lstm \_layer1\_25 | 0 | 0.0002 | 0 | 10 | 50 | 0.00005 | 512 | 25 | 0 | 64 | 92% | 25 |
| model\_second\_dense\_layer\_ dropout\_L2 | 0 | 0.0002 | 0.2 | 10 | 50 | 0.00005 | 512 | 50 | 0 | 64,32 | 35% | 25 |
| model\_second\_dense\_layer\_ nodrop | 0 | 0.0002 | 0 | 10 | 50 | 0.00005 | 512 | 50 | 0 | 64,32 | 90% | 25 |
| model\_second\_lstm\_layer\_no \_dropout\_L2.ipynb | 0 | 0.0002 | 0 | 10 | 50 | 0.00005 | 256 | 50 | 25 | 64 | 86% | 30 |
| model\_second\_lstm\_layer\_no \_dropout\_L2\_30\_history\_points | 0 | 0.0002 | 0 | 30 | 80 | 0.00005 | 512 | 50 | 25 | 64 | 97% | 35 |
| model\_no\_dropout\_L2\_50\_history\_ points | 0 | 0.0002 | 0 | 50 | 80 | 0.00003 | 200 | 50 | 0 | 64 | 96% | 30 |
| model\_no\_dense\_layer\_1\_no\_ dropout\_L2 | 0 | 0.0002 | 0 | 10 | 80 | 0.00005 | 512 | 50 | 0 | 0 | 96% | 30 |
| model\_L1\_Regulation\_nodropout \_no\_dense\_layer1 | 0.0002 | 0 | 0 | 10 | 50 | 0.00005 | 512 | 10 | 0 | 0 | 98% | 50 |
| model\_25\_history\_point\_test\_no\_ dense\_L1\_no\_dropout\_L2 | 0 | 0.0002 | 0 | 25 | 50 | 0.00003 | 512 | 50 | 0 | 0 | 97% | 25 |
| model\_no\_dense\_layer\_1\_no\_ dropout\_L1\_Regulation\_50\_history\_ points\_100\_epochs | 0.0002 | 0 | 0 | 50 | 100 | 0.00005 | 512 | 50 | 0 | 0 | 98% | 15 |
| model\_no\_dense\_L1\_no\_ dropout\_50\_history\_point\_ test\_batchsize\_120\_ L2\_Regulataion | 0 | 0.0002 | 0 | 50 | 50 | 0.00003 | 120 | 50 | 0 | 0 | 95% | 45 |

סדר הסקריפטים בטבלה הינו מייצגת את סדר העבודה הכרונולוגי .

\*\* זמן הריצה נגזר מריצה של המחשב שלי כאשר רצו עליו מספר תוכניות במקביל. ייתכנו זמנים אחרים בשימוש בפלטפורמות אחרות

תיאור מהלך העבודה:

לאחר שהתחלנו עם מודל הבסיס ניסינו להוסיף לו רגולציות, אך ראינו שזה לא מועיל, לכן אמרנו בוא ננסה להוריד את הDropOut ואכן חל שיפור משמעותי החל מהסקריפט הרביעי. עשינו גם ניסוי לבדוק מה יקרה עם נפעיל את שני הרגולציות, ידענו שזה לא יעבוד כמו שצריך כיוון שהמודל בקושי יבצע שינויים כי יש לו מיסוי כפול על כל צעד ,ואכן כך גילינו. ביצענו גם שינוי של ה batch size לקטן יותר אך לא ראינו שיפור. מה שיכולנו ללמוד מזה , הוא שהדאטה שלנו מספיק גדול כך שכדאי להשאיר את batch size על גודל עצום של 512 . מעבר לזה זה גורם לתנודתיות בתוצאות, ככל הנראה בגלל העדכון התכוף למשקלים. התחלנו להבחין כי המודל עם רגולציה של L1 מוביל על פני ה L2 ,אך בכל זאת המשכנו לבחון אותו כדי לראות אם הוא מתחיל להיות טוב יותר עם פרמטרים אחרים. שמנו לב כי ככל שנוריד את הlearning rate יקח יותר זמן למודל להגיע לבערך אותה תוצאה,דיי צפוי, לכן אין צורך יותר מידי לגעת בו . ניסינו גם לשנות את הגודל של המודל LSTM גם כלפי מעלה וגם כלפי מטה, אך מה שקרה זה חוסר יציבות וקפיצות מאוד חדות בגרף. ניסנו שתי ניסיונות עם שתי שכבות רגילות (dense layer ), פעם אחת עם dropout (שהיה צפוי שיהיה די גרוע) ועוד ניסיון ללא dropout וראינו שוב פעם את ההפרש המטורף עם הdropout וללא ,הפרש של קרוב ל55 אחוז. מה שחתם בעבורנו את האיסור להשתמש בdropout .  
ביצענו עוד ניסוי להוסיף עוד רשת LSTM אולי היא תשפר דבר או מה , אך לא ראינו שיפור משמעותי. ואז עלה לנו רעיון לנסות להגדיל פשוט את הכמות איברים שהוא מסתמך עליהם, ראינו בלא מעט מקומות שנותנים יותר מידע אחורה לכן החלטנו לנסות, העלנו מ10 ל30 , על אותו הסקריפט וגילינו שיפור משמעותי והיינו מרוצים והמשכנו לנסות את זה על בגדול יותר על המודל הרגיל וגם ראינו תוצאות טובות. לאחר מכן , ניסינו לשנות את צורת המודל עוד פעם ולהוריד את השכבות הרגילות בכלל, והשינוי היה לטובה גם הוא . הוא הצליח להוסיף עוד איזה אחוז או שתיים למודלים . ולבסוף עשינו ניסוי של מספר שינויים יחדיו , אך הוא היה פחות מוצלח.

בתחילת הדרך בדקנו גם את הגודל של הרגולציות L2 ו L1 אך לא תיעדנו זאת לצערנו, ומפאת הזמן אין באפשרותנו לעלות את הניסויים בסקריפטים.