**בגרסה ראשונה**

מכניסים גדלים שונים של זיכרון כלומר מספר שונה של מדידות אחורה.

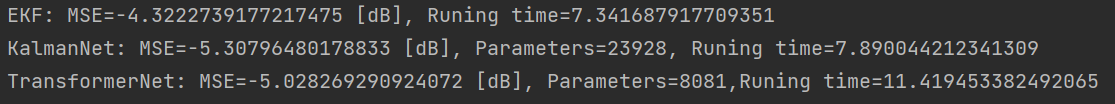
ביציאה מהטרנספורמר מבצעים Padding ככה שטנזור היציאה יהיה בעל גודל קבוע

וממשיכים עם FC בגודל קבוע.

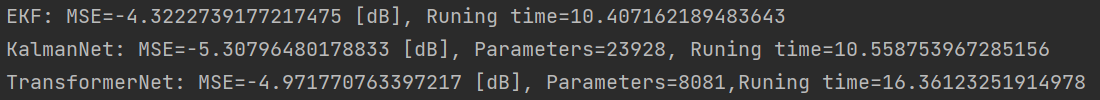
ללא שימוש ב MASK

מאומן לקבלת התוצאות הטובות ביותר עבור זיכרון של 6 צעדים אחורה בשלושת התוצאות מטה.

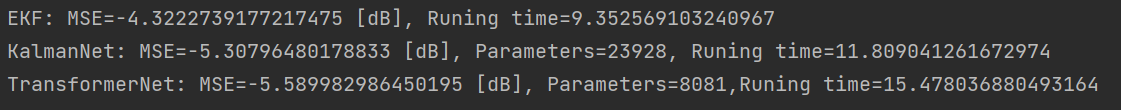
באימון בכל באצ' הוגדר גודל שונה לזיכרון בין 1 ל 12



באימון בכל באצ' הוגדר גודל שונה לזיכרון בין 4 ל 8

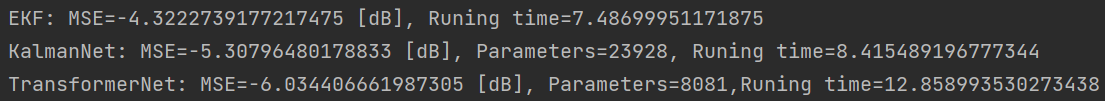


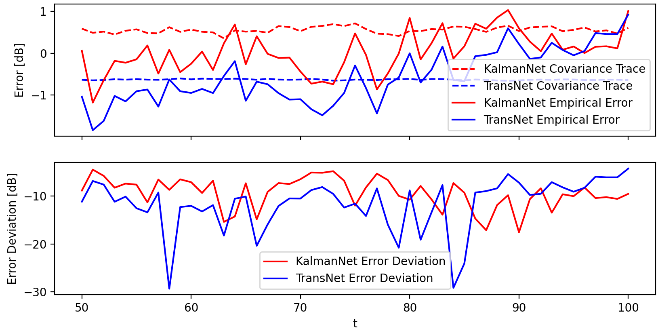
באימון בכל באצ' הוגדר גודל שונה לזיכרון קבוע 6



באימון בכל באצ' הוגדר גודל שונה לזיכרון קבוע 12

נראה יציב יותר מהקודם. EMPERICAL ERROR נמוך מKN





**גרסה שניה.**

1. השינוי, נבצע פאדינג לפני הכניסה לטרנספורמר.( יעיל יותר מבחינת ביצועים על GPU )
2. בחירת מודל עם שינוי מינימאלי עבור גודל שונה של זיכרון ושמירה על ביצועיים מייטביים
3. יצירת "ראש" נוסף לשיערוך מיטוב הביצועים עבור הגדלה/הקטנה של הזיכרון

~~בדיקה של שימוש בMASK והזנה של הערכים הנכונים GROUND TRUTH~~

בדיקה של הרצה מספר פעמים בלולאה בפנים DEAFUTION (מעלה זמן ריצה)

שינוי פרמטרים של רשת הטרנספורמרים וההשפעה שלהם

הגדלת הזיכרון מאד מאד FNET פורייר טרנספורם ובדיקת שינוי על הביצועים ביחס לדרישות המערכת ( זמן ריצה, מספר פעולות)

**גרסה 3**

ביצוע שיוך של מדידות למסלול (שימוש ברשת שנוצרה כבר..)

***שאלות:***

האם מודל תנועה נכון יותר משפיע על ביצועי הרשת?

האם הכנסה של P משפר את ביצועי הרשת?

דיוק של המודל עם הזמן, נראה שהמודלים מתכנסים אחד לשני לאחר זמן מסוים? האם נכון?