

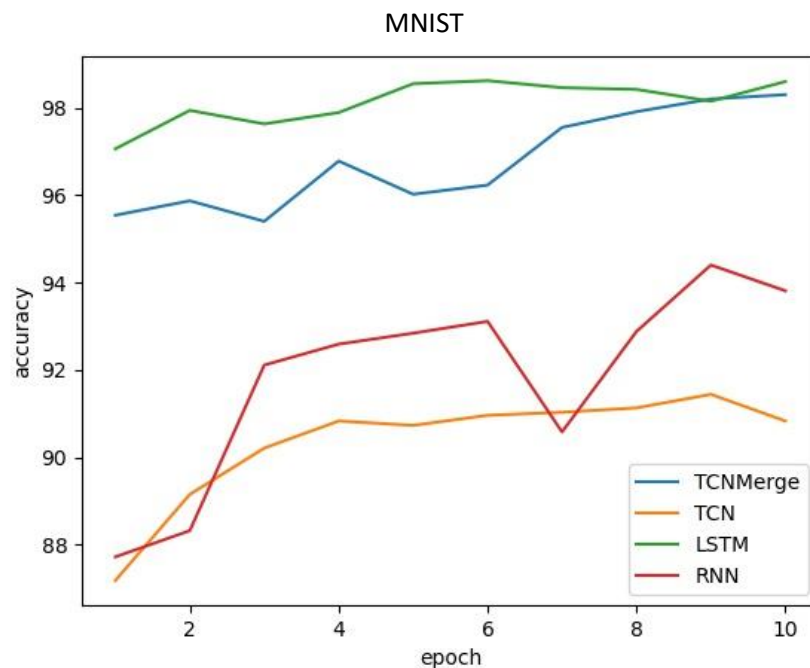
נושאים בלמידה עמוקה - סיכום וניסיון שחזור תוצאות מאמר

An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling/Shaojie Bai, J. Zico Kolter, Vladlen Koltun

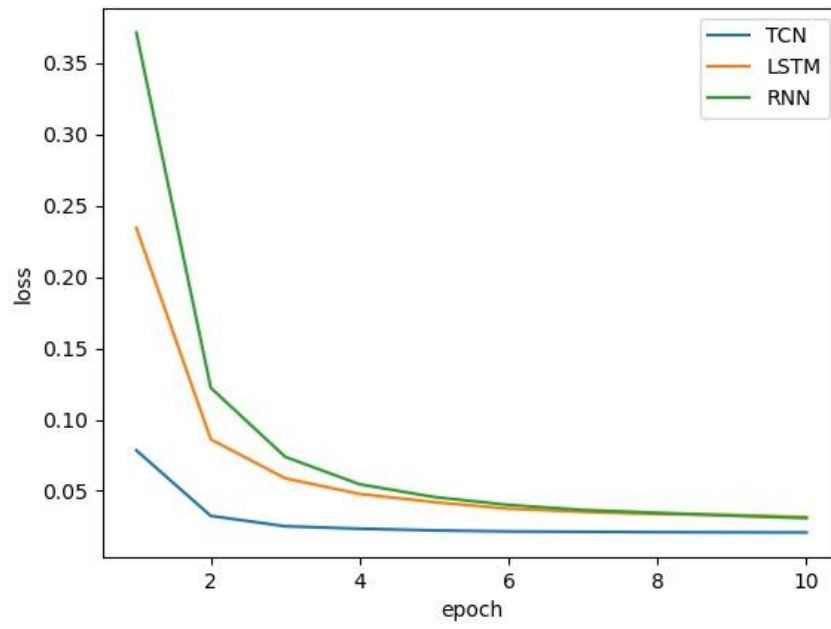
במאמר זה ניסו החוקרים לאתגר את האסוציאציה השכיחה בין משימות סדרתיות לבין RNN. החוקרים בנו רשת קונבולוציונית גנרית ופשוטה יחסית, והשוו בינה לבין Recurrent Networks כמו LSTM, RNN ו GRU, במספר רב של משימות שברובן נהוג להשתמש ב RNN על כלל גזרותיה. החוקרים קראו לרשת Temporal Convolutional Networks, או בקיצור TCN, והיא מורכבת מ Residual Blocks, אשר מורכבים מסדרה של קונבולוציות מורחבות, נרמול, פונקציות אקטיבציה ו dropouts. תוצאות המחקר מראות כי ה TCN עלתה בביצועים על ה Recurrent Networks, מה שגרם לחוקרים להעריך כי הזכרון של רשת ה TCN ארוך יותר מהזכרון של הרשתות שמולן היא התחרתה, ולכן יתכן כי היא נקודת פתיחה טובה יותר עבור משימות סדרתיות הדורשות זכרון רב מאשר RNN.

בפרויקט זה ניסינו לשחזר את תוצאות הניסוי של החוקרים. השווינו ב 3 משימות בין רשת TCN המורכבת מ Residual Block יחיד לבין RNN ו LSTM, ובנוסף ניסינו להרחיב את TCN ל TCNMerge אשר מחלקת את הקלט לשניים, מעבירה כל חלק ב Residual Block, מאחדת את התוצאה ומעבירה אותה ב RB נוסף. להלן תוצאות הניסוי:

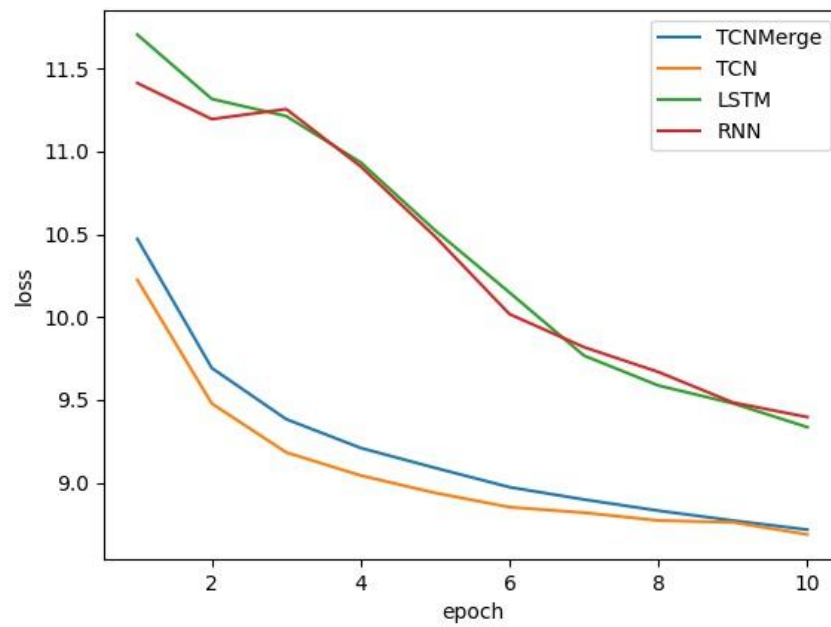
| TCNMerge | TCN | LSTM | RNN | |
|----------|---------|---------|---------|---------------------------|
| 98.41% | 91.21% | 98.15% | 91.72% | MNIST(accuracy) |
| - | 0.02060 | 0.02868 | 0.03262 | COPY TASK(loss) |
| 8.71642 | 8.68761 | 9.33522 | 9.39618 | MUSIC – JSB(loss) |
| 3.70297 | 3.94192 | 4.52659 | 4.04973 | MUSIC – Nott(loss) |



COPY TASK (T=100)



MUSIC - JSB



MUSIC - Nott

