LSTMs Exploit Linguistic Attributes of Data

Nelson F. Liu, Omer Levy, Roy Schwartz, Chenhao Tan, Noah A. Smith
LSTMs work well for natural language data

Are they particularly well-suited for language?
Testbed Memorization Task

• Given a constant-length sequence of \( k \) inputs, recall the identity of the middle token.

• Task is inherently **non-linguistic**, inputs can be arbitrary sequences.
Linguistic Data Improves Memorization Performance
Linguistic Data Improves Memorization Performance
Linguistic Data Improves Memorization Performance
Linguistic Data Improves Memorization Performance

![Graph showing the effect of different linguistic data on accuracy with varying input sequence lengths.](image-url)
So, are LSTMs particularly well-suited for language?

Yes, more than uniform data or data with selected linguistic attributes
LSTMs solve the task by counting
More Questions

• How does the LSTM use linguistic patterns in training?

• What happens when you add more hidden units?

if you want to know more...come to our poster!