



► **RESEARCH STAY WEEK 3, LANGUAGE MODELING AND N-GRAMS**

Daniel Cajas A01708637

► CONTEXT

Historically, n-grams have been a big force behind NLP. There have been a plethora of techniques but n-grams have proved a tried and true technique that is computationally cheap compared to modern neural network based models. However it does have its limitations as the context window grows exponentially and the models tend to be more short sighted.

► **SEARCH METHODOLOGY**

The material given focuses on n-grams, their construction, technical challenges and how they have advanced. I wanted to see how they fit in the ecosystem of NLP as of today. That is why, through google scholar, I looked for recent efforts on how to leverage the power of n-grams by improving them in other models.

► COMPARISON

As Erick Bril, et al. n.d. mentions. N-grams have given the best results for many NLP tasks over the years despite their lack of linguistic information. Over the years hardware progress, model development and infrastructure work have made it possible to build even larger and more powerful n-grams which have proved useful. They have the advantage of having a small cost for lookup as its essentially just a dictionary lookup, instead of intensive matrix multiplications like with neural networks.

If that is the case, why not combine them and get the best of both worlds? Aurko Roy, et al. 2022 did just that. They created a neural network layer based on n-grams as the entry tokens of the network. With this change they found that a smaller model could perform almost equally to the bigger models, at a considerably reduced computational cost.