**Meeting Notes: April 18 2021**

<https://pytorch.org/hub/pytorch_fairseq_translation/>

Dana Found A GOOD TRANSLATION

<https://he.wikisource.org/wiki/%D7%9E%D7%A7%D7%A8%D7%90_%D7%9E%D7%91%D7%95%D7%90%D7%A8/%D7%AA%D7%A6%D7%95%D7%92%D7%94>

Convert Dataset to No nekud and compare to text with Nekud.

(Mention that in the discussion section.)

TODO:

* Dataset Creation
  + Nekud
  + No nekud
  + Structure:
    - Start with a CSV.
    - Each column is a pasuk.
    - Then python from there to tokenize stuff.
    - Check how the dataset looks in pset3
    - Check how the dataset looks in PyTorch Pretrained models [here.](https://pytorch.org/hub/pytorch_fairseq_translation/)

| Old Hebrew | Modern Hebrew | Pusak | Perek | Url to page |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |

* Models:
  + **"Vanilla" Seq-to-seq**
    - PSET 3 (or maybe find an implementation with **attention**).
    - "Some things to remember writing about is that the order of words is not consistent".
  + **The one Jacob sent: (Zachary)**
    - <https://arxiv.org/pdf/1704.04368.pdf>
    - implementations
      * <https://github.com/abisee/pointer-generator>
      * <https://github.com/atulkum/pointer_summarizer>
      * https://github.com/jiminsun/pointer-generator
      * https://opennmt.net/
  + **PyTorch Transformer Implementation (Dana)** 
    - <https://pytorch.org/hub/pytorch_fairseq_translation/>
    - <https://pytorch.org/hub/huggingface_pytorch-transformers/>
    - https://pytorch.org/tutorials/intermediate/seq2seq\_translation\_tutorial.html

<https://he.wikisource.org/wiki/%D7%9E%D7%A7%D7%A8%D7%90_%D7%9E%D7%91%D7%95%D7%90%D7%A8#%D7%AA%D7%95%D7%A8%D7%94>

1. Professor recommended:
   1. <https://arxiv.org/pdf/1704.04368.pdf>
   2. implementations
   3. <https://github.com/abisee/pointer-generator>
   4. <https://github.com/atulkum/pointer_summarizer>
2. Translations currently: <https://pytorch.org/hub/pytorch_fairseq_translation/>
   1. <https://pytorch.org/hub/huggingface_pytorch-transformers/>

THREE

One from the pset

FOUR

Look to middle english to modern english--models they use for this.

<https://github.com/tokestermw/tensorflow-shakespeare>   
<https://github.com/ashleynicole472/Shakespeare-Translator>  
<https://github.com/parthpatel20010/Shakespeare-to-Modern-English-Translator>

FIVE

TA sent prepublished paper

<https://github.com/jlrussin/syntactic_attention>

Necud vs no necud, nikud.

Model from the pset.