Deep Spell

Next Generation Query Auto-Suggestion in NDS

FTS Suggestion now

- Suggestion based on Inverted Token Index

Term	Occurences	Prio
10th	10.000	?
new	1000	1
york	50	1
main	800	10
st	100.000	?
…	…	…

- Problem: Context has no influence on suggestion

St L[ouis]	vs	St Loui[siana]	→ Subtoken violation
California Los A[ngeles]	vs	California Los A[lamos]	→ Hierarchy violation
Virginia V[ictoria]	vs	Virginia V[irginia]	→ Redundancy violation

Solution

- (This is how Apple does it!)
- Leave query string processing to Neural Network (NN):
  - Tokenization
    - NN input is full user string, not just last term:
      - NN Input: “New York C”
      - Old FTS Input: “C”
  - Categorization
    - NN categorizing hierarchy classes before completion, so actually:
      - User Input: “California Palo A”
      - NN Input: “California Palo A” → [“California”]$_{\text{STATE}}$ [“Palo A”]$_{\text{CITY}}$
  - Suggestion
    - NN returns probability distribution for following characters:
      - suggest([“California”]$_{\text{STATE}}$ [“Palo A”]$_{\text{CITY}}$) = $\left(\begin{bmatrix} l & p=0.3 \\ n & p=0.1 \\ \dots & \dots \end{bmatrix}\begin{bmatrix} t & p=0.5 \\ a & p=0.1 \\ \dots & \dots \end{bmatrix}\begin{bmatrix} o & p=0.3 \\ e & p=0.2 \\ \dots & \dots \end{bmatrix}\right)$

Architecture