Systems Analysis
Santa 2024

Gabriela Martínez Silva

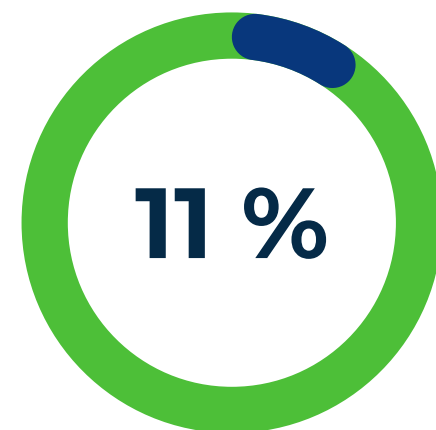**Jairo Arturo Barrera Mosquera**

Main points

- **INPUT**:Shuffuled words

- **TASK**:Reconstruct the sequence into a passage

Evaluation

01 Evaluation model

For each sequence in subbmission the perplexity is caculated usiing GAMMA 2 9B

02 Perplexity

11 %

It essentially quantifies the uncertainty or "surprise" a model has when predicting the next item (such as a word) in a sequence

Data Strcture

Input

ID   **TEXT**

1   **advent chimney elf family fireplace gingerbread mistletoe ornament reindeer scrooge**

Output

ID   **TEXT**

1   **elf family fireplace advent chimney gingerbread mistletoe ornament reindeer scrooge**

Example*

sequence	
id 🔑	integer
text	string

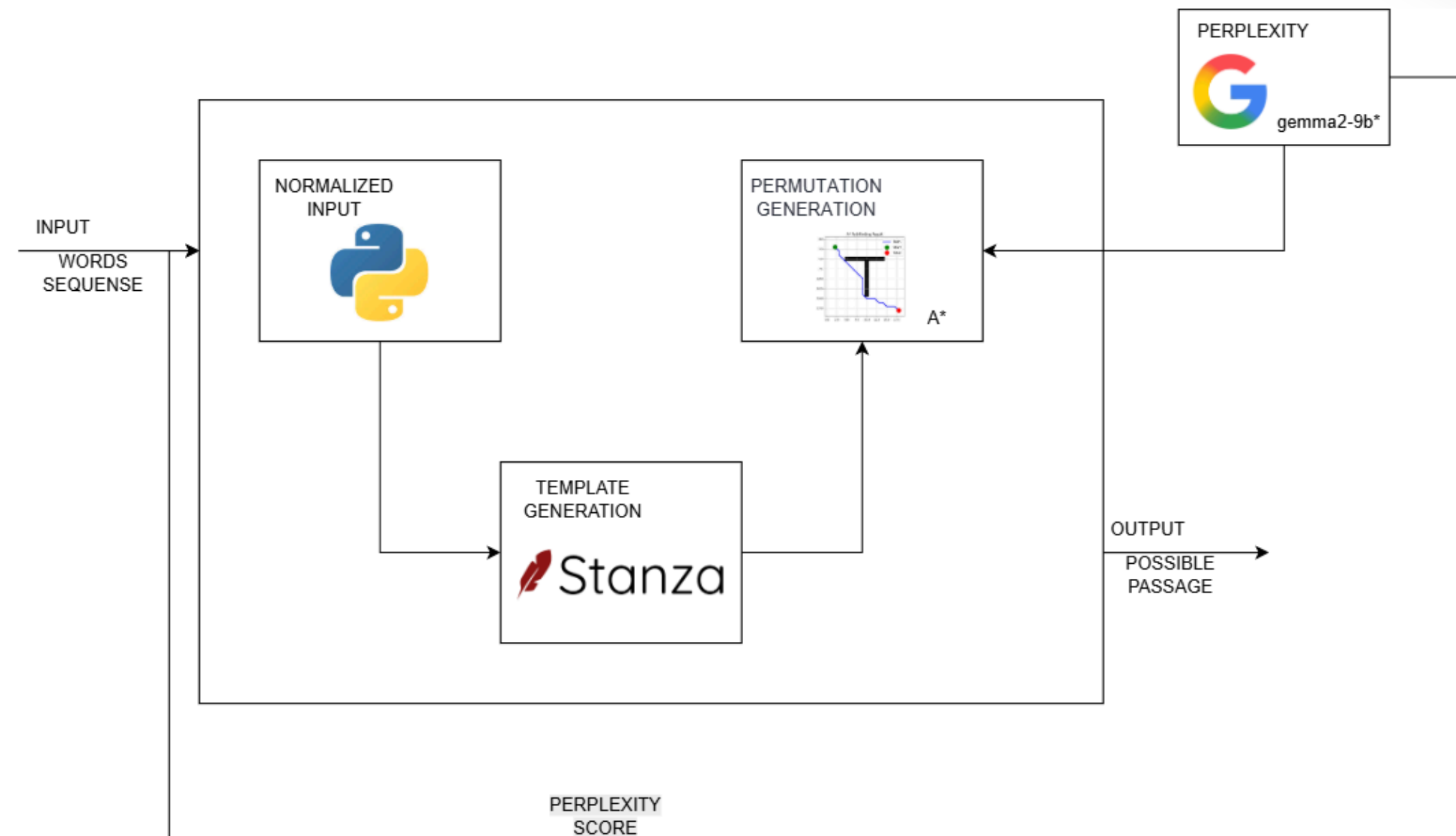# System requirements

Functional

- Sequence Length Handling

- Word Constraints

- Sentence Usage

Non-Functional

▲ Perplexity

▲ Semantic Coherence:

Our Solution

Systematic approach for solve the problem

Components

Normalized Input

- Reading data
- Formating data

Define Templates

Propose one or more structural "templates" that a valid sentence permutation should ideally follow.

Generate Permutations

Using the templates present in the second element of the systems generate possible solutions to the reordering problem

01

02

03

Normalize input

Objective

Prepare the raw data

Functions

Reading, Validation, Basic Cleaning, Word Count.

Output

Structured word list

Templates generation

Objective

Infer grammatical structure

Functions

Linguistic Analysis (POS tagging), for templates generation

Output

Linguistic features and potential templates/constraints.

Permutation genration

Objective

Create candidate sentences

Functions

Genrate final passages using Intelligent Search Algorithms (A*)

Output

Partial Passage and final ppassages

Evaluation(Feedback)

Objective

Rate the coherence/fluency of a candidate permutation.

Feedback

Perplexity of a partial candidate permutation

Sensitivity

- Text Length: The greater the number of words in a sentence, the more possible permutations exist, making optimization more difficult

- Grammatical Structure: Some passages may contain ambiguities that impact the perplexity evaluation.

- Lexical distribution: Phrases with common words can generate more plausible permutations, while phrases with rare words may have fewer viable options

Conclusions

- Complex problem, systematic solution

- Linguistic analysis and intelligent search.

- Cybernetic approach

Thanks

GITHUB

🌐 https://github.com/vydibot/Systems-Analysis