



Mehrzaad Ahmadian

M.Sc. in Information Technology

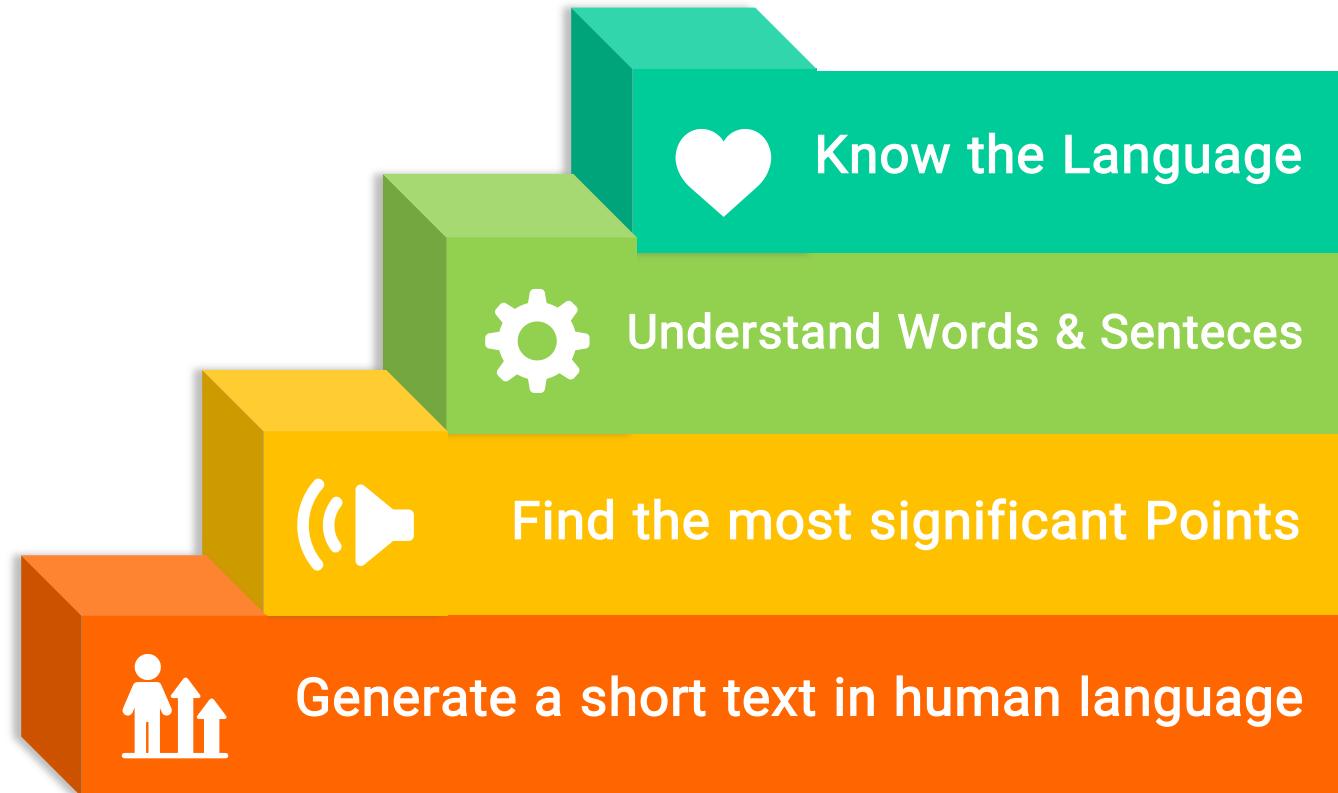
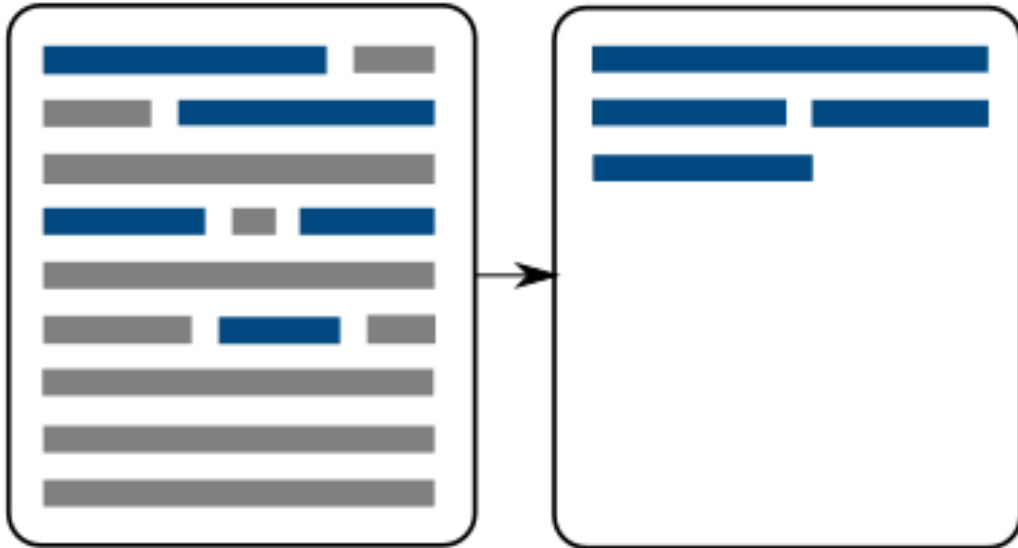
Thesis title:

**Interactive query-based summarization of scientific
documents considering qualitative feedback**

Text Summarization



Generate a short text containing the most significant points in the original text

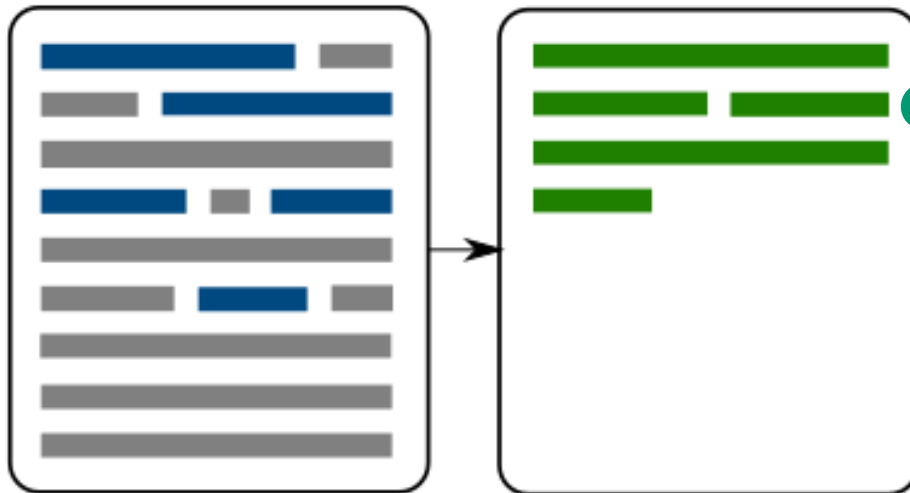
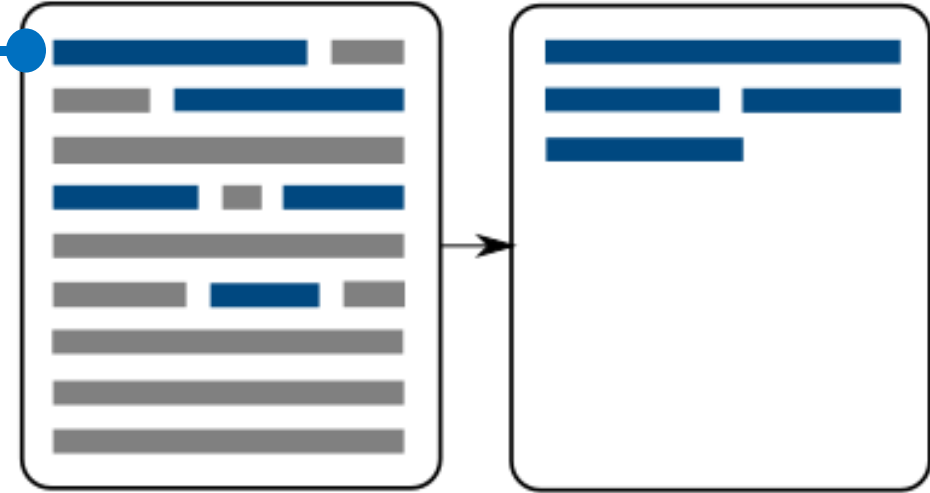


Extractive / Abstractive Summarization



Extractive

- Sentences Feature Extraction
- Scoring base on features
- Sorting
- Concatenation



Abstractive

- Preprocessing
- Encode the original text
- Autoregressive summary generation

Generic / Query-Based



G

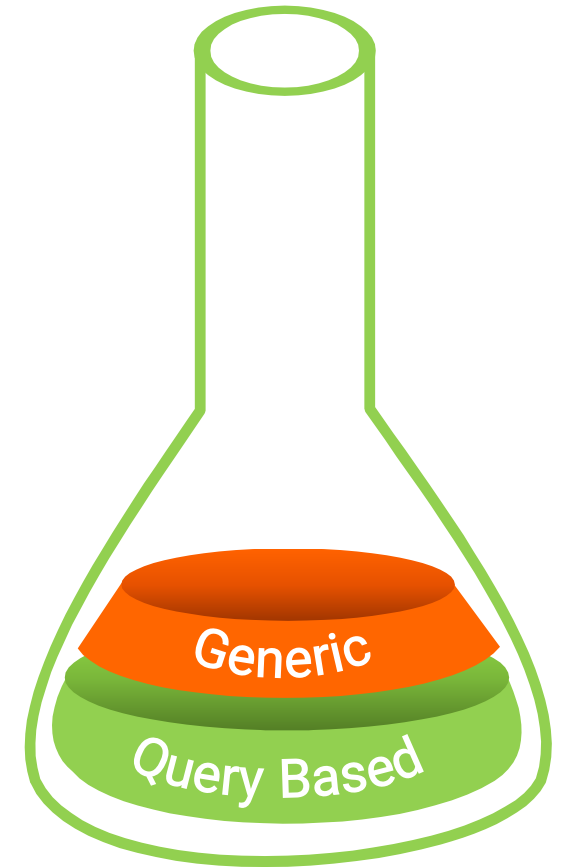
Generic

- Generate based on the most important sentences
- Without considering the user need

Q

Query-Based

- Generate based on the text and user query
- Considering the user need



Related works (Persian Summarization)



Abstractive



Query-Based



New approach to overcome lack of query-based summarization dataset in the Persian language



Considering User Feedback



Interactive query-based summary generator

Based on Transformer Architecture...

Summarization type:

☒ Query-Based ☐ Generic

Query:

سهم بنزین مصرفی منطقه را چه کسی اعلام کرد؟

Translation of the Query:

"Who announced the amount of gasoline consumed in the region?"

Text to summarize:

مصرف نفت کوره منطقه سبزوار در بهار امسال، نسبت به مدت مشابه پارسال، 58 درصد کاهش یافت. مدیر شرکت ملی پخش فراورده های نفتی منطقه سبزوار به خبرنگار شانا، گفت: مصرف نفت کوره این منطقه به دلیل توسعه شبکه گاز طبیعی در سه ماه نخست امسال، با 58 درصد کاهش نسبت به مدت مشابه سال پیش، به 370 هزار لیتر رسیده است. عیسی افتخاری سهم بنزین مصرفی این منطقه را نیز در بازه زمانی یاد شده 38 میلیون و 975 هزار لیتر اعلام و اظهار کرد: این مقدار از افزایش چهار درصدی مصرف این سوخت در قیاس با بهار پارسال حکایت دارد. به گفته وی، در 93 روز نخست امسال، 46 میلیون و 236 هزار لیتر نفت گاز در منطقه سبزوار مصرف شده که نسبت به مدت مشابه سال پیش، با افزایش نیم درصدی همراه شده است. مدیر شرکت ملی پخش فراورده های نفتی منطقه سبزوار همچنین، از مصرف سه میلیون و 959 هزار لیتر نفت سفید در این منطقه خبر داد و افزود: مصرف این سوخت در قیاس با سه ماه نخست پارسال، 26 درصد کاهش

Generate summary

Clean text

Generated summary: (28 words and 131 characters)

مدیر شرکت ملی پخش فراورده های نفتی منطقه سبزوار، از مصرف ۲۸ میلیون و ۹۷۵ هزار لیتر بنزین در سه ماه نخست امسال در این منطقه خبر داد.

Rate the relevance of the generated summary to your query:



Other suggested queries you may want to use:

سهم بنزین مصرفی منطقه را چه کسی اعلام کرد؟

مصرف نفت کوره سبزوار در چه سالی به چه صورت رسیده است؟

نفت سفید در سبزوار در چه دوره کاهش یافته است؟

مصرف گاز طبیعی فشرده در سبزوار در چه سالی بود؟

نفت گاز در سبزوار در چه دوره ای مصرف شده است؟

Translation of the Summary:

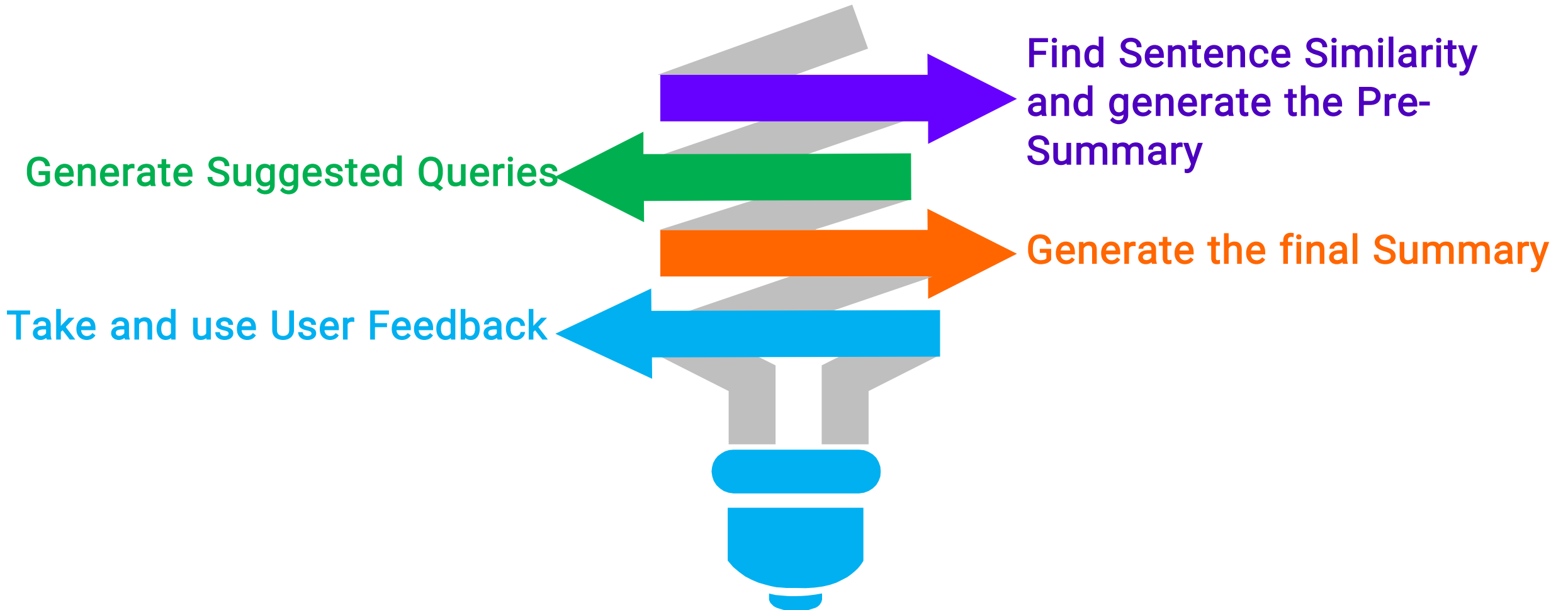
"The director of the National Oil Distribution Company in Sabzevar region announced the consumption of 38 million and 975 thousand liters of gasoline in the first three months of this year in this region."

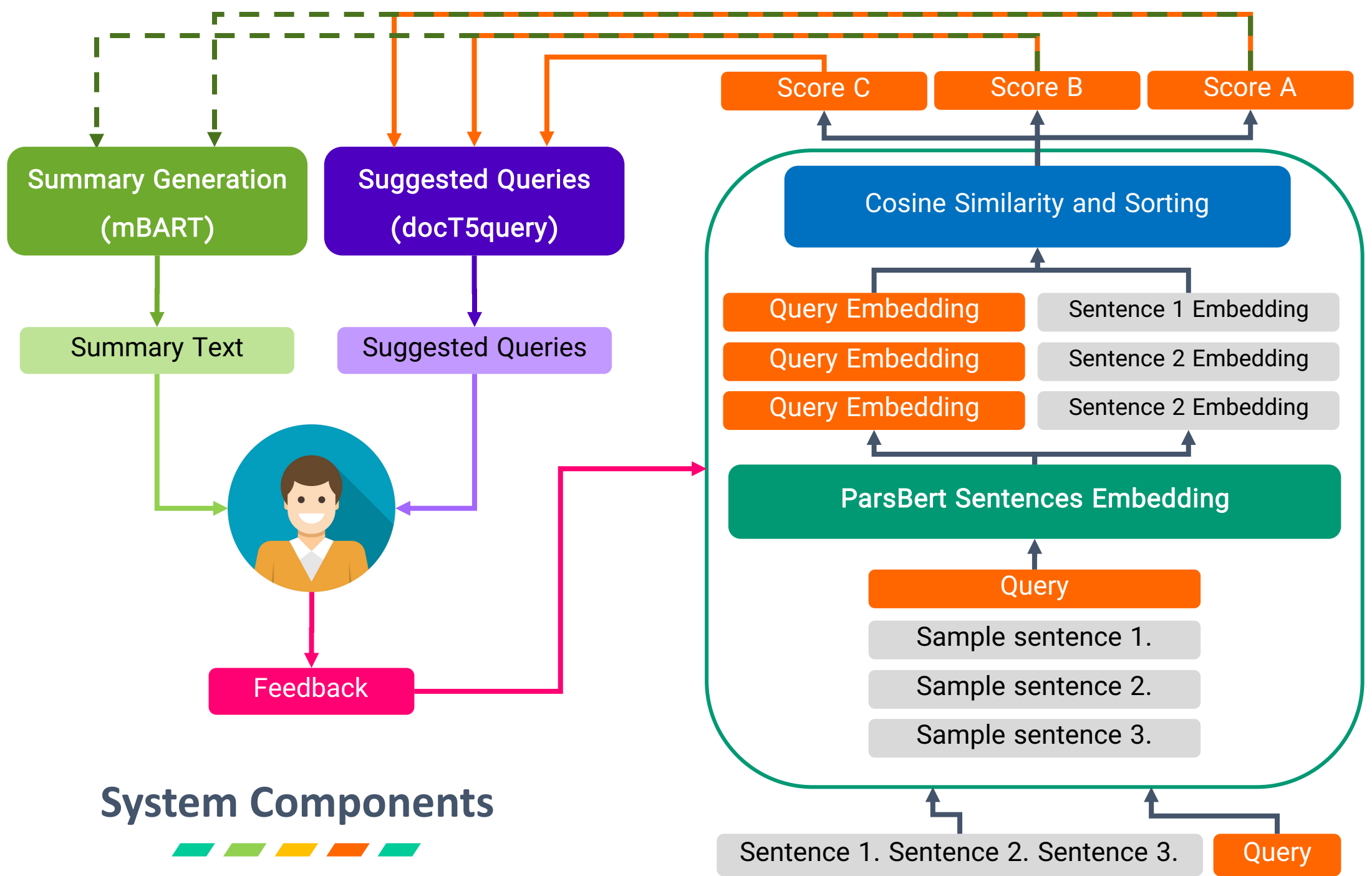
Final Deployed Application



<https://ahmadian.pro/summarizer>

System Parts



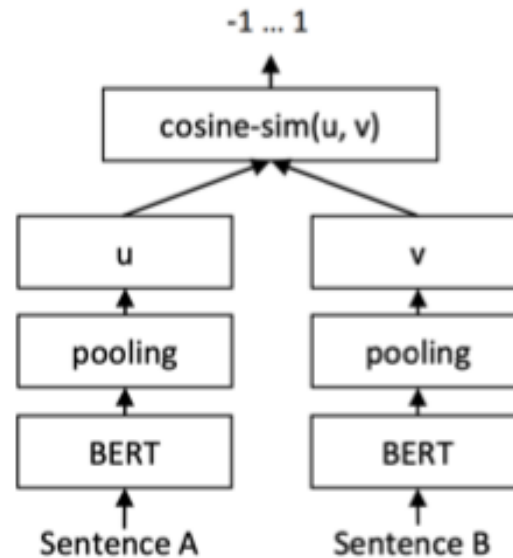


System Components



Sentence Similarity and Pre-Summary Generation

- Used SBERT (ParsBERT) to calculate sentences embeddings.
- Fine-tuned on FarsTail Dataset
- Sort and generate Pre-Summary



System Components (Contd.)



Final Summary Generation

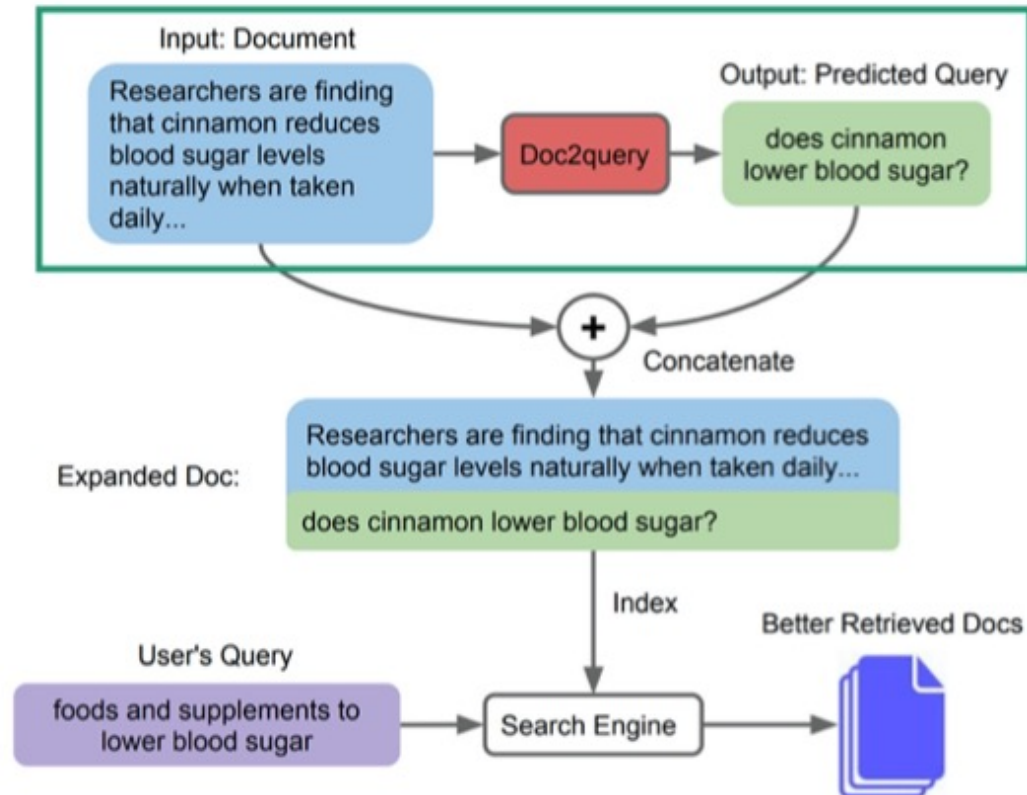
- Used mBART-50.
- Pre-trained on 53 languages by facebook.
- Fine-tuned on pn-summary dataset for Persian generic summarization.

System Components (Contd.)

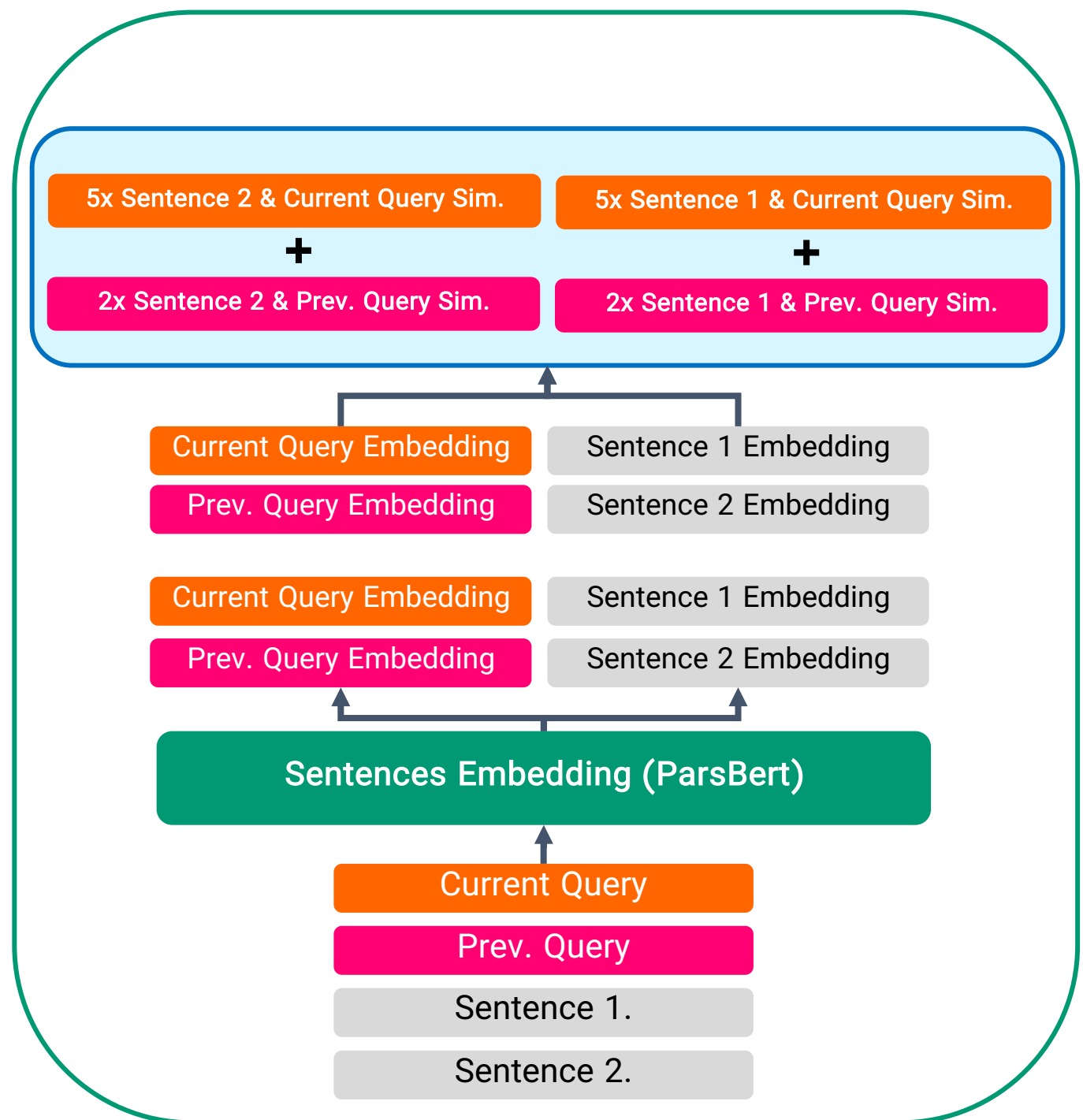


Suggested queries generation

- Used mT5-base model.
- Pre-Trained on C4 dataset by google.
- Fine-tuned on PersianQA dataset (for QA task).



Feedback System





Interactive query-based summary generator

Based on Transformer Architecture...

Summarization type:

☒ Query-Based ☐ Generic

Query:

سهم بنزین مصرفی منطقه را چه کسی اعلام کرد؟

Translation of the Query:

"Who announced the amount of gasoline consumed in the region?"

Text to summarize:

مصرف نفت کوره منطقه سبزوار در بهار امسال، نسبت به مدت مشابه پارسال، 58 درصد کاهش یافت. مدیر شرکت ملی پخش فرآورده های نفتی منطقه سبزوار به خبرنگار شانا، گفت: مصرف نفت کوره این منطقه به دلیل توسعه شبکه گاز طبیعی در سه ماه نخست امسال، با 58 درصد کاهش نسبت به مدت مشابه سال پیش، به 370 هزار لیتر رسیده است. عیسی افتخاری سهم بنزین مصرفی این منطقه را نیز در بازه زمانی یاد شده 38 میلیون و 975 هزار لیتر اعلام و اظهار کرد: این مقدار از افزایش چهار درصدی مصرف این سوخت در قیاس با بهار پارسال حکایت دارد. به گفته وی، در 93 روز نخست امسال، 46 میلیون و 236 هزار لیتر نفت گاز در منطقه سبزوار مصرف شده که نسبت به مدت مشابه سال پیش، با افزایش نیم درصدی همراه شده است. مدیر شرکت ملی پخش فرآورده های نفتی منطقه سبزوار همچنین، از مصرف سه میلیون و 959 هزار لیتر نفت سفید در این منطقه خبر داد و افزود: مصرف این سوخت در قیاس با سه ماه نخست پارسال، 26 درصد کاهش

Generate summary

Clean text

Generated summary: (28 words and 131 characters)

مدیر شرکت ملی پخش فرآورده های نفتی منطقه سبزوار، از مصرف ۲۸ میلیون و ۹۷۵ هزار لیتر بنزین در سه ماه نخست امسال در این منطقه خبر داد.

Rate the relevance of the generated summary to your query:



Other suggested queries you may want to use:

سهم بنزین مصرفی منطقه را چه کسی اعلام کرد؟

مصرف نفت کوره سبزوار در چه سالی به چه صورت رسیده است؟

نفت سفید در سبزوار در چه دوره کاهش یافته است؟

مصرف گاز طبیعی فشرده در سبزوار در چه سالی بود؟

نفت گاز در سبزوار در چه دوره ای مصرف شده است؟

Translation of the Summary:

"The director of the National Oil Distribution Company in Sabzevar region announced the consumption of 38 million and 975 thousand liters of gasoline in the first three months of this year in this region."

Final Deployed Application

<https://ahmadian.pro/summarizer>

Summarization type:

☒ Query-Based ☐ Generic

Query:

مصرف نفت سفید چقدر کاهش داشته است؟

Translation of the Query:

"How much has the consumption of kerosene decreased?"

Text to summarize:

مصرف نفت کوره منطقه سبزوار در بهار امسال، نسبت به مدت مشابه پارسال، 58 درصد کاهش یافت. مدیر شرکت ملی پخش فرآورده های نفتی منطقه سبزوار به خبرنگار شانا، گفت: مصرف نفت کوره این منطقه به دلیل توسعه شبکه گاز طبیعی در سه ماه نخست امسال، با 58 درصد کاهش نسبت به مدت مشابه سال پیش، به 370 هزار لیتر رسیده است. عیسی افتخاری سهم بنزین مصرفی این منطقه را نیز در بازه زمانی یاد شده 38 میلیون و 975 هزار لیتر اعلام و اظهار کرد: این مقدار از افزایش چهار درصدی مصرف این سوخت در قیاس با بهار پارسال حکایت دارد. به گفته وی، در 93 روز نخست امسال، 46 میلیون و 236 هزار لیتر نفت گاز در منطقه سبزوار مصرف شده که نسبت به مدت مشابه سال پیش، با افزایش نیم درصدی همراه شده است. مدیر شرکت ملی پخش فرآورده های نفتی منطقه سبزوار همچنین، از مصرف سه میلیون و 959 هزار لیتر نفت سفید در این منطقه خبر داد و افزود: مصرف این سوخت در قیاس با سه ماه نخست پارسال، 26 درصد کاهش

Generate summary

Clean text

Generated summary: (19 words and 93 characters)

مصرف نفت سفید منطقه سبزوار در سه ماه نخست امسال، نسبت به مدت مشابه پارسال، ۲۶ درصد کاهش یافت.

Rate the relevance of the generated summary to your query:



Other suggested queries you may want to use:

نفت سفید در سبزوار در چه دوره کاهش یافته است؟

سهم بنزین مصرفی منطقه را چه کسی اعلام کرد؟

مصرف نفت کوره سبزوار در چه سالی به چه صورت رسیده است؟

نفت گاز در سبزوار در چه دوره ای مصرف شده است؟

تعداد فرآورده نفتی سبزوار در نخستین فصل امسال به چه تعداد رسیده است؟

Translation of the Summary:

"Kerosene consumption in Sabzevar region decreased by 26% in the first three months of this year compared to the same period last year."

Query records:

Row	Score	Query
1	☆☆☆☆☆	سهم بنزین مصرفی منطقه را چه کسی اعلام کرد؟

Final Deployed Application

<https://ahmadian.pro/summarizer>



Thanks for your attention



Mehrzaad Ahmadian