Trade news summarisation

Lyutova Daria

May 2024

Abstract

This research aims to enhance the effectiveness of summarization in the context of trade news analysis. Generating summaries for trade news articles poses a significant challenge due to the specific demands and constraints of the task. The summarization process necessitates careful consideration of data requirements, including the inclusion of essential elements such as country references, product details, and key figures from the news. This study concentrates on data exploration and the implementation of a simplistic model designed to process portions of the text efficiently. Also Large Language Model (LLM) approaches were tried. The results were evaluation using metrics such as BERT-score.

Project code: `https://github.com/lyutovad/trade_news_summary`

1 Introduction

The project involving the summarization of trade news for the portal is of significant importance due to the growing volume of information in the trade industry. With an abundance of news articles published daily in multiple languages, it becomes crucial to extract key insights efficiently to keep up with the fast-paced trade environment. By summarizing trade news, stakeholders can quickly obtain essential information such as import/export data, company mentions, and country/product references. This task not only saves time but also enhances decision-making processes, enabling users to stay informed about the latest trends and developments in the global market. Therefore, this project plays a vital role in facilitating access to pertinent trade information and streamlining the analysis of trade-related news for professionals and enthusiasts in the field.

In the context of news articles on trade, there are two distinct categories or levels: simple and complex news. Simple news articles typically focus on a limited number of products and countries, making it easier to extract key information for summarization. These articles are characterized by clear and concise references to specific products and countries, facilitating the summarization process.

On the other hand, complex news articles pose a challenge for summarization due to the scattered distribution of information regarding products and countries throughout the text. In complex news articles, the products and countries mentioned are not concentrated in specific sections but are rather interspersed throughout the entire text. When creating summaries for complex news articles manually, annotators often face the task of gathering all references to products and countries from various parts of the text before condensing the information into a coherent summary. This process adds complexity and requires careful attention to detail to ensure the summary accurately captures the essence of the original article.

2 Related Work

Text summarization involves two primary approaches: extractive and abstractive. Extractive methods select the most relevant phrases or sentences from the original text to create a summary, while abstractive methods rephrase the text, potentially altering the original semantic content.

When utilizing the extractive approach, one of the main difficulties lies in determining which sentences are most crucial to include in the final summary. With a vast amount of text to sift through, it can be challenging to identify the most relevant and important information that captures the essence of the original content. This process requires careful consideration and analysis to ensure that the summary accurately represents the key points of the text while maintaining coherence and readability. Additionally, the subjective nature of determining importance can further complicate the task, making it essential for individuals to have a clear understanding of the content and its objectives.

There are several articles that offer different methods for selecting sentences to create a text summary. Graph-based biased LexRank approach combined with topic modeling for topic-based extractive summarization model [K Zheltova, 2022], useing of pretrained language models like BERT to rank sentences and create a concise summary which can extract aspects [Anonymous, 2021].

MemSum uses a complex Markov decision process with history awareness to create effective summaries. It combines information from sentence encoders, global context, and past extractions to produce concise summaries that outperform other models [Nianlong Gu, 2022].

The vectors on which models are trained are essential for any approach. For example, [Yang Liu, 2019] introduce a document-level encoder based on BERT for semantic expression and sentence representation. Their extractive model uses this encoder with Transformer layers. For abstractive summarization, authors recommend a fine-tuning schedule with separate optimizers for the encoder and decoder to tackle the pretraining mismatch. A two-stage fine-tuning approach further enhances summary quality. Good pretrained sentence representations are not only necessary for supervized learning, but as shown by [Qianren Mao, 2023] the Utilizing pre-trained embeddings optimized for cohesive and distinctive sentence representations improves sentence ranking in extractive summarization.They propose a novel graph pre-training method that models intra-sentential features and inter-sentential cohesion through sentence-word bipartite graphs. These embeddings enhance unsupervised summarization performance as well.

As for abstractive summarization the development of advanced techniques or this approach was significantly influenced by the introduction of the Transformer model in the article "Attention Is All You Need." [Ashish Vaswani, 2017] This model's innovative self-attention mechanism improved the quality of summarization by capturing long-range dependencies and contextual information. Many pretrained models have emerged for the sammarization task such as a PEGASUS [Jingqing Zhang, 2020] where essential sentences are excluded or masked from the input document and then created as a single output sequence from the remaining sentences, resembling an extractive summary. Or LOCOST [Florian Le Bronnec, 2024] with it's architecture of encoder-decoder models based on state-space models for conditional text generation with lengthy contextual inputs.

A promising approach for this task involves combining an extractive method with subsequent processing using Transformer models with automated post-editing. For instance, in compressive-based summarization, we can extract important sentences from a news article based on named entities identified in the previous step, and then compress them. To facilitate compression, solutions like the Syntactically Look Ahead Attention Network [Hidetaka Kamigaito, 2020] offer a Seq2Seq network with lookahead syntactic attention, enabling the generation of informative summaries by tracking parent-child word dependencies during decoding and capturing crucial words for future decoding. The repository includes all scripts for extracting embeddings, fine-tuning models, and making predictions on the Google dataset. However, one of the steps in the setup process takes nearly 1 day and requires 300GB of disk space.

3 Model Description

For our task, we chose the MT5 (Multilingual Translation 5) model [Linting Xue, 2021], which is a highly promising choice for text summarization tasks involving two languages and potentially more languages in the future. Its versatility lies in its ability to handle multilingual text data effectively. This model can be particularly beneficial for summarization tasks due to its capacity to understand and generate content in multiple languages, which can enhance the quality and accuracy of the summaries. As more languages are added to the MT5 model, it is expected to offer even greater flexibility and performance when summarizing texts in diverse language pairs.

Also the TinyLlama-1.1B-Chat model [Peiyuan Zhang, 2024] was selected for experimentation based on its impressive characteristics and suitability for different tasks and following the instructions. With a relatively compact size compared to larger language models and rapid inference speed, it is well-suited for efficient processing in chat applications. This Transformer-based model, boasting 1.1 billion parameters, a vocabulary size of 32000, and a context length of 2048, was fine-tuned on the UltraChat dataset containing synthetic dialogues generated by ChatGPT. Additionally, FlashAttention mechanism, alignment with the TRL DPOTrainer using the openbmb/UltraFeedback dataset, comprising 64,000 clues and model enhancements rated as "GPT-4," further enhances its capabilities.

The third model - Mistral Instruct v0.2 [Albert Q. Jiang, 2023] was selected for summarization due to its tailored features for instruction-following tasks and thematic coherence maintenance. With a 7 billion parameter size and Transformer architecture similar to GPT-3, this model excels in executing instructions and adherence to specified topics. Trained on the Pile dataset, Mistral Instruct utilizes special tokens for instruction separation and features a larger context window in v0.2. The model emphasizes language understanding and efficiency enhancements, outperforming other models in various benchmarks, making it a promising choice for summarization tasks.

4 Dataset

The dataset for the trade news portal consists of news articles related to trade topics in both Russian and English languages. The final output of the dataset comprises concise digests of the news articles. Each digest is required to include digital indicators of imports/exports, other product-related metrics, mentions of companies in the news, countries referenced in the news, and products mentioned in the news. Additionally, the dataset preprocessing involves removing introductory words and quotes, replacing vague time references with specific dates obtained from the news article itself, and ensuring company mentions are included in the summary. All summaries in the dataset were manually curated by annotators following the technical specifications provided by the customer. On the Tab. 1 you can see the statistics for the mentioned dataset.

	Russian	English
Articles	26,921	
Texts	19,543	7,378
Text Lemmas	86,703	86,377
Text Words	157,096	90,864
Summary Lemmas	34,409	22,054
Summary Words	62,406	23,496

Table 1: Dataset statistics of the news dataset. The numbers represent the quantity of articles, texts in Russian and English, lemma and word counts for both texts and summaries.

4.1 Data exploration

Data exploration and analysis were conducted to enhance the model performance by investigating the impact of the quantity of products and countries on the length of the summaries, as well as determining the extent to which all specified products and locations are covered within the summaries. Surprisingly, the quantity of products and countries did not directly influence the length of the summaries. In some cases, the length of the summaries depended on factors beyond the mention of locations or products.

The length of news articles often increases due to the need to provide context, such as in articles discussing the lifting of sanctions where a reminder of the history of sanctions is essential. This relevant context may exist in the text but may not be obvious during the summarization process.

It was observed that locations have a more significant influence on the summary text than the quantity of products. When numerous countries are mentioned in the text, their list is condensed into a union including all mentioned countries. Similarly, a large quantity of products can be simplified into the general Standard International Trade Classification (SITC) code. However, the classification of products by the SITC code faces challenges due to the classifier's low quality, making it unreliable for text processing.

Also in English articles, when a product is classified under code 93 (special operations and products not classified by type), it may not be immediately clear which product is being referenced. Therefore, explanations or descriptions of such products are necessary in the summaries to ensure comprehension, especially for Russian readers who may not be familiar with them.

Summaries containing quantitative data were investigated, with entities such as money, percent, quantity, price, and currency identified. Among these entities, only a total of 254 news articles contained all these entities. Currency, deemed to lack crucial information, was excluded from the analysis.

After consolidating all the data with entities into a single dataset, there were 18,907 rows remaining (13,531 in Russian and 5,376 in English). While no linear relationship was found between the length of the summary and numerical indicators, a slight correlation was observed for the quantity entity with the summary length in the dataset containing all entities.

The analysis of numerical indicators' occurrence in summaries revealed that a significant portion of numerical data was included in the summaries, but establishing a clear dependency between the number of matches and the length of the text or summary length was challenging.

A hypothesis was proposed to rank entities by importance in the selection process based on named entity recognition (NER) lists, prioritizing locations with numerical values over products. Furthermore, in cases where multiple numerical indicators were identified, a specific sequence for their selection was suggested: quantity, percent, price, and money.

5 Experiments

5.1 Metrics

For all annotated summaries, the BERTScore metric [Tianyi Zhang and Artzi, 2020] with model type "bert-base-multilingual-cased" was evaluated. BERTScore is commonly used in text generation tasks due to its superior correlation with human ratings compared to traditional overlap-based methods. The advantages of using this metric for summarization models include:

- Contextual embeddings, such as those from BERT and similar models, provide a better understanding of the text's meaning, which is crucial for identifying semantic similarity between the original text and its concise summary. This leads to a more precise evaluation of summarization quality.
- Cosine similarity between contextual embeddings helps assess how closely the summary content aligns with the original text, providing insight into the substantive similarity between the texts. Traditional word overlap-based metrics cannot offer this level of understanding.
- Linear scaling of metric values facilitates the interpretation of evaluation results, aiding in faster and more accurate model adjustments for summarization tasks. An additional benefit of the metric for our task may be the consideration of rare words encountered in news texts.

BERTScore allows for the examination of three metrics: precision, recall, and F1 score.

$$R_{\mathrm{BERT}} = \frac{1}{|x|} \sum_{x_i \in x} \max_{\hat{x}_j \in \hat{x}} \mathbf{x}_i^\top \hat{\mathbf{x}}_j \ , \quad P_{\mathrm{BERT}} = \frac{1}{|\hat{x}|} \sum_{\hat{x}_j \in \hat{x}} \max_{x_i \in x} \mathbf{x}_i^\top \hat{\mathbf{x}}_j \ , \quad F_{\mathrm{BERT}} = 2\frac{P_{\mathrm{BERT}} \cdot R_{\mathrm{BERT}}}{P_{\mathrm{BERT}} + R_{\mathrm{BERT}}} \ .$$

Figure 1: Recall, precision, and F1 scores.

- Recall BERTScore reflects how comprehensively the summarization model conveys information from the source text. It measures the number of tokens in the reference sentence that match tokens in the candidate sentence. Higher recall BERTScore values indicate more comprehensive information transmission by the model.
- Precision BERTScore indicates how accurately the summarization model conveys information from the source text. It measures the number of tokens in the candidate sentence that match tokens in the reference sentence. Higher precision BERTScore values indicate more accurate information transmission by the model.

- F1 BERTScore combines precision and recall BERTScore values, representing the overall performance of the summarization model. It accounts for both the accuracy and completeness of information transfer from the source text. Higher F1 BERTScore values indicate more precise and comprehensive information transmission by the model.

It is important for us to focus more on the recall metric or the F1 score. Additionally, considering the F-beta score, where beta approaches 2, can be beneficial in evaluating summarization model performance.

5.2 Experiment Setup

Two models of the mT5 multilingual family were trained, and the model was manually selected by the annotators based on the summary quality of 299 texts. The data was split into 80/20 proportions for training and testing. Table 2 shows the metrics of the models with different number of training epochs and different hyperparameters. Although the best metrics were those of the mt5 model with one iteration of pretraining, the annotators found the summaries generated by the model that was pretrained for 2 epochs to be better - shorter, but still containing more information than the other models generated.

	precision	recall	f1
mt5_2it_en_ru	0.848253	0.62331	0.717865
mt5_1it_en_ru	**0.86779**	**0.664231**	**0.751126**
mt5_4it_en_ru	0.861589	0.657617	0.744693
mt5-small-1it-en-ru	0.858532	0.652348	0.740165
mt5-small-2it-en-ru	0.865639	0.663803	0.750058
mt5-small-4it-en-ru	0.854428	0.649016	0.736534
mt5_2_2it_en_ru	0.847128	0.623615	0.717717
mt5_2_4it_en_ru	0.817424	0.609012	0.697118

Table 2: bert-score metrics for digests generated by 2 models with different sets of parameters. (299 texts).

At the moment, the csebuetnlp/mT5_multilingual_XLSum model was added to the news processing pipeline. The training involved 26921 news articles with handwritten summaries according to the customer's TOR. 19543 Russian texts and 7378 English texts. The model successfully processes simple news with a small number of products and countries. However, for complex news (examples in Appendix A) the model does not work properly and the digest needs to be manually adjusted. Table 3 presents the bert-score metrics for manual summaries, the chosen model, and random text mixtures (equivalent to the baseline) across the entire dataset.

The figure 2 depicts the distribution of the metric BERT-score recall for 3 models. Config of the model is in section B

	manual	mT5_multilingual_XLSum	random
mt5_2it_en_ru	0.848253	0.62331	0.717865
precision	0.855738	0.848048	0.630689
recall	0.685404	0.628855	0.547413
f1	0.759816	0.721564	0.585786
std_precision	0.052763	0.031111	0.030102
std_recall	0.081000	0.048699	0.028576
std_f1	0.065605	0.039727	0.026016

Table 3: bert-score metrics for manual summaries, the chosen model, and random text mixtures (all data)

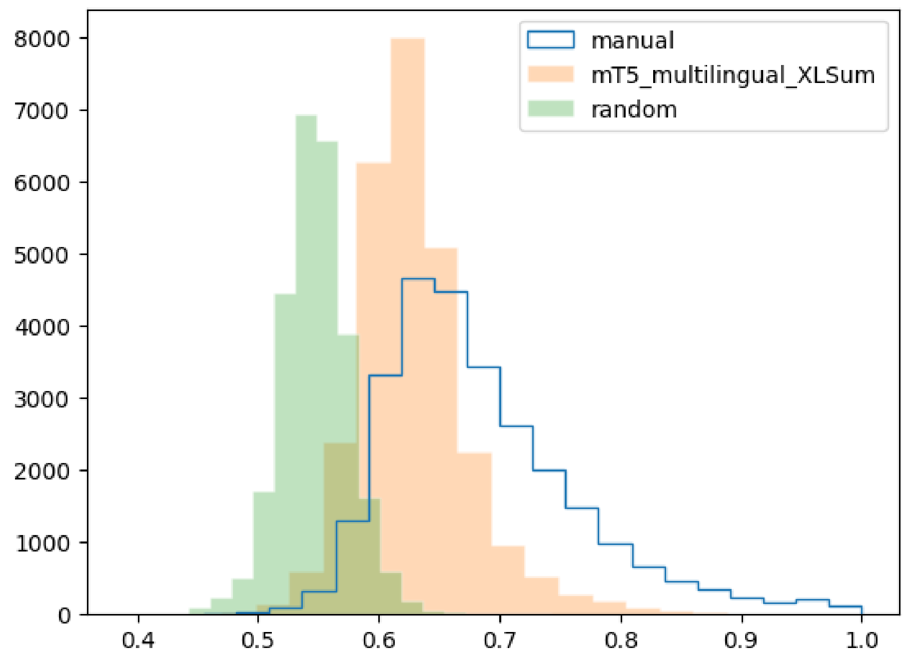

Figure 2: BERT-score recall for manual summaries, the chosen model, and random text mixtures (all data)

When it comes to large language models, they were chosen to assess training speed, inference speed, and quality of summarization.

The model was trained on 1500 texts, 10000 texts, and the entire dataset. Hyperparameters were modified twice, and 3 different prompts were utilized. The results are presented in the table 4.

The best prompt: Summarize the following text. Focus on countries, products and numbers. If goods can be grouped in one group with the same name,

	precision	recall	f1
1500 texts, hyperparameters1, prompt1	0.732	0.624	0.672
10 000 texts, hyperparameters2, prompt2	**0.746**	**0.645**	**0.691**
all data, hyperparameters1. prompt3	0.684	0.634	0.657

Table 4: tinyllama results

it is necessary to do so. Don't include names and introductory words. be impersonal and use bullet points.

Mean inference speed: 12 s.

Although the bert-score recall is higher for the best bert-score model than for the simple model, the quality of the summary is very poor: the summary is poorly aligned with the text. When pretrained on an incomplete dataset, noise generation was observed.

Hallucinations are present in the generated summary.

In addition, for the model trained on all data, all news items were translated into English.

Regarding Mistral Instruct v0.2, it was also trained using LoRA with the best prompt. However, the inference time for the model ranges from 12 to 13 minutes per news article. Therefore, validation on the dataset comprising 299 news articles was not conducted, as the model is not suitable for practical use.

An example of a generated summary for a complex news article is provided in the appendix C.

Mt5: did not convey the necessary indicators correctly.

TinyLlama: provided excessive details on insignificant information.

Mistral: mixed up a location.

5.3 Baselines

As a baseline on this dataset, an extractive model has long been utilized, selecting the first 5 sentences mentioning countries. However, the quality of these digests is poor, with metrics aligning with random scoring. As a result, all summaries were manually written due to the inadequacy of the automated approach.

6 Results

The main outcome of this stage of the project is a functional model that still requires adjustments but demonstrates significantly improved performance compared to the previous extractive approach. This model is currently replacing part of the annotator's tasks and streamlining work processes.

However, the search for suitable solutions is still ongoing. Currently, the following approach appears promising for experiments:

1. Utilize entities identified by NER in the previous step: countries, products, numerical indicators.
2. Extract sentences containing these entities.
3. Aggregate entities into summaries. [Clément Jumel, 2020], [Jose Angel, 2022]
4. Compress sentences. [Hidetaka Kamigaito, 2020]
5. Verify factual consistency at the entity level. [Feng Nan, 2021]

7 Conclusion

At this research stage, an analysis of previous studies was carried out. Summaries incorporating locations, products, and numerical data were scrutinized. Various Transformer models were tested to select the most suitable one, which was then integrated into the news processing pipeline and is now actively utilized. Furthermore, the performance quality and speed of the Large Language Models were thoroughly assessed. The following stages of the project have been assessed.

References

[Albert Q. Jiang, 2023] Albert Q. Jiang, Alexandre Sablayrolles, A. M. C. B. D. S. C. D. d. l. C. F. B. G. L. G. L. L. S. L. R. L. M.-A. L. P. S. T. L. S. T. L. T. W. T. L. W. E. S. (2023). Mistral 7b. *Computation and Language.*

[Anonymous, 2021] Anonymous (2021). Extractive topical summarization with aspects. *ACL ARR.*

[Ashish Vaswani, 2017] Ashish Vaswani, Noam Shazeer, N. P. J. U. L. J. A. N. G. L. K. I. P. (2017). Attention is all you need. *NeurIPS.*

[Clément Jumel, 2020] Clément Jumel, Annie Louis, J. C. K. C. (2020). Tesa: A task in entity semantic aggregation for abstractive summarization. *EMNLP.*

[Feng Nan, 2021] Feng Nan, Ramesh Nallapati, Z. W. C. N. d. S. H. Z. D. Z. K. M. B. X. (2021). Entity-level factual consistency of abstractive text summarization. *arXiv.*

[Florian Le Bronnec, 2024] Florian Le Bronnec, Song Duong, M. R. A. A. N. F. C. V. G. A. L. L. S. P. G. (2024). Locost: State-space models for long document abstractive summarization. *paperswithcode.*

[Hidetaka Kamigaito, 2020] Hidetaka Kamigaito, M. O. (2020). Syntactically look-ahead attention network for sentence compression. *paperswithcode.*

[Jingqing Zhang, 2020] Jingqing Zhang, Yao Zhao, M. S. P. J. L. (2020). Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. *ICML*.

[Jose Angel, 2022] Jose Angel, Annie Louis, J. C. K. C. (2022). Source-summary entity aggregation in abstractive summarization. *COLING*.

[K Zheltova, 2022] K Zheltova, A Ianina, V. M. (2022). Topical extractive summarization. *Conference on Artificial Intelligence and Natural Language*, pages 14–26.

[Linting Xue, 2021] Linting Xue, Noah Constant, A. R. M. K. R. A.-R. A. S. A. B. C. R. (2021). mt5: A massively multilingual pre-trained text-to-text transformer. *NAACL*.

[Nianlong Gu, 2022] Nianlong Gu, Elliott Ash, R. H. R. H. (2022). Memsum: Extractive summarization of long documents using multi-step episodic markov decision processes. *ACL*.

[Peiyuan Zhang, 2024] Peiyuan Zhang, Guangtao Zeng, T. W. W. L. (2024). Tinyllama: An open-source small language model. *Computation and Language*.

[Qianren Mao, 2023] Qianren Mao, Shaobo Zhao, J. L. X. G. S. H.-B. L. J. L. (2023). Bipartite graph pre-training for unsupervised extractive summarization with graph convolutional auto-encoders. *arXiv*.

[Tianyi Zhang and Artzi, 2020] Tianyi Zhang, Varsha Kishore, F. W. K. Q. W. and Artzi, Y. (2020). Bertscore: evaluating text generation with bert. *ICLR 2020*.

[Yang Liu, 2019] Yang Liu, M. L. (2019). Text summarization with pretrained encoders. *IJCNL*.

A Examples of complex news and manual summaries

Text1

США: заявки на стальной импорт в январе упали на 6% к предыдущему месяцу Как сообщает Американский институт железа и стали (AISI), количество заявок на получение разрешений на импорт стали за январь составило 2,359,000 нетто-тонн. Это на 6,0% меньше по сравнению с 2,509,000 разрешенных тонн, зафиксированными в декабре, и на 12,7% больше, чем в декабре, когда общий объем окончательного импорта составил 2,093,000 тонн. Тоннаж разрешений на импорт готовой стали в январе составил 1,777,000 тонн, что на 9,5% больше общего объема конечного импорта в 1,622,000 в декабре. Предполагаемая доля рынка импорта готовой стали в январе составила 21Импорт стали со значительным увеличением разрешений в январе по сравнению с окончательным импортом в декабре включает листы и полосы со всеми другими металлическими покрытиями (рост на 145%), катанку (рост на 58%), механические трубы (рост на 39%), горячекатаные прутки (рост на 39 процентов) и проволока тянутая (до 26%). В январе наибольшее количество заявок на получение разрешений на импорт стали поступило в Канаду (599 000 тонн, рост на 11% по сравнению с декабрьским итогом), Бразилию (446 000 тонн, рост на 33%), Мексику (286 000 тонн, снижение на 2%), Южную Корею (130 000 тонн, снижение на 52%) и Японии (110 000 тонн, рост на 41%).

Summary1

Количество заявок на получение разрешений на импорт стали за январь 2024 года составило 2,359,000 нетто-тонн. Это на 6,0% меньше по сравнению с 2,509,000 разрешенных тонн, зафиксированными в декабре 2023-го. Тоннаж разрешений на импорт готовой стали в январе составил 1,777,000 тонн, что на 9,5% больше общего объема конечного импорта в 1,622,000 в декабре. Предполагаемая доля рынка импорта готовой стали в январе составила 21%. В январе 2024 года наибольшее количество заявок на получение разрешений на импорт стали поступило в Канаду (599 000 тонн, рост на 11% по сравнению с декабрьским итогом), Бразилию (446 000 тонн, рост на 33%), Мексику (286 000 тонн, снижение на 2%), Южную Корею (130 000 тонн, снижение на 52%) и Японии (110 000 тонн, рост на 41%).

Text2

Поставки СПГ из России во Францию выросли на 41 процент МИД: поставки СПГ из России во Францию за 9 месяцев 2023 года выросли на 41% "По итогам 2022 года Россия стала вторым после США поставщиком природного газа во Францию за счет импорта российского СПГ. В январе-сентябре 2023 года французские закупки из России превысили аналогичный период 2022 года на 41%", — сказал Студенников. Он добавил, что основным пунктом приема для российских поставок сжиженного газа стал терминал в Монтуар-де-Бретань, в который в 2022 году поступило 1,68 миллиарда кубических метров, а в первом полугодии 2023 года — 0,89 миллиардов кубических метров СПГ. Ранее газета Telegraph со ссылкой на данные Евростата сообщала, что страны ЕС в 2023 году закупили более половины экспортируемого Россией СПГ, при этом Испания и Франция стали вторым и третьим крупнейшими покупателями российского СПГ после Китая. Согласно данным статистической службы ЕС, Испания закупила российского СПГ на 1,8 миллиарда евро в первые девять месяцев 2023 года. Франция заняла второе место в ЕС по этому показателю, закупив российского СПГ на 1,5 миллиарда евро. Следующая за ней Бельгия потратила на те же цели 1,36 миллиарда евро. Европейский рынок был исторически главным для российских газовиков, поставки туда газа в прошлом достигали 150 миллиардов кубометров и выше. Однако в

2022 году — после физического уничтожения "Северных потоков", а также санкций против альтернативных маршрутов поставок — экспорт из России существенно снизился. Поставки газа "Газпрома" в дальнее зарубежье по итогам 2022 года упали примерно на 85 миллиардов кубометров, до 100,9 миллиарда. Практически все это падение пришлось на ЕС.

Summary2

По итогам 2022 года Россия стала вторым после США поставщиком природного газа во Францию за счет импорта российского СПГ. В январе-сентябре 2023 года французские закупки из России превысили аналогичный период 2022 года на 41%.

Text3

Число дилеров LCV в России за год снизилось почти на 400 На конец января 2024 года в нашей стране числится 1115 официальных дилерских центров* по продаже и обслуживанию легких коммерческих автомобилей LCV. Такие данные получили эксперты агентства «АВТОСТАТ» в ходе подготовки обзора «Дилерские сети. Мониторинг дилерских контрактов по легким коммерческим автомобилям». Эксперты отмечают, что за год этот показатель снизился на 388. Напомним, что в январе прошлого года на территории РФ насчитывалось 1503 дилера LCV. Общее снижение количества дилерских центров связано с тем, что в течение года было расторгнуто 524 действующих контракта, а новых заключено – 136. Многие «закрывшиеся» за это время дилеры LCV теперь предоставляют только услуги сервиса. Больше всего таковых оказалось у компании Renault – вследствие ухода с российского рынка все 120 дилеров французской марки перешли в разряд сервисных партнеров. На 108 дилеров меньше стало у Hyundai (96 из них выполняют только функции сервиса), на 47 – у Isuzu, на 32 – у Volkswagen, на 26 – у Peugeot, на 25 – у Mercedes-Benz, на 21 – у Citroen, на 14 – у Opel. При этом даже на фоне такой негативной тенденции есть и такие компании, у которых с начала нынешнего года число дилеров LCV увеличилось. Прежде всего, это относится к китайской марке Foton – за год у нее стало на 44 дилерских центра больше. У недавнего новичка рынка – отечественного Sollers – общее количество дилеров выросло на 13, а у ГАЗа и АВТОВАЗа – на 8 и 3 соответственно. Эксперты агентства «АВТОСТАТ» считают, что в ближайшие месяцы число дилерских центров LCV будет расти, т.к. ожидается приход новых игроков в этот сегмент рынка, прежде всего, из Китая (Forland, JMC и другие). В этой связи отметим и отечественный КАМАЗ, запустивший продажи модели «Компас», которая также имеет легкую коммерческую версию. А некоторые производители LCV, уже работающие на нашем рынке, проводят перезапуск дилерских сетей, например, Dongfeng и Foton. Более подробная информация о дилерских центрах LCV (с их контактными данными) в разбивке по брендам и регионам содержится здесь. Источник: https://www.autostat.ru/news/56828/ © Автостат.

Summary3

На конец января 2024 года в России числится 1115 официальных дилерских центров (без учета отдельных сервисных центров) по продаже и обслуживанию легких коммерческих автомобилей LCV. За год этот показатель снизился на 388. В январе 2023 года на территории РФ насчитывалось 1503 дилера LCV. Общее снижение количества дилерских центров связано с тем, что в течение года было расторгнуто 524 действующих контракта, а новых заключено – 136. Многие «закрывшиеся» за это время дилеры LCV теперь предоставляют только услуги сервиса. Больше всего таковых оказалось у компании Renault – вследствие ухода с российского рынка все 120 дилеров французской марки перешли в разряд сервисных партнеров. На 108 дилеров меньше стало у Hyundai (96 из них выполняют только функции сервиса), на 47 – у Isuzu, на 32 – у Volkswagen, на 26 – у Peugeot, на 25 – у Mercedes-Benz, на 21 – у Citroen, на 14 – у Opel. При этом даже на фоне такой негативной тенденции есть и такие компании, у которых с начала 2024 года число дилеров LCV увеличилось. Прежде всего, это относится к китайской марке Foton – за год у нее стало на 44 дилерских центра больше.

Text4

Vit Nam's exports surge 42% in January February 10, 2024 - 14:29 Vit Nam's exports reached nearly US$33.6 billion in January, a 42 per cent surge over the same period last year and the highest level since April 2022. HCM CITY — Vit Nam's exports reached nearly US$33.6 billion in January, a 42 per cent surge over the same period last year and the highest level since April 2022. According to a report by the Ministry of Industry and Trade, Vit Nam's total imports and exports of goods in the first month surpassed $64 billion, up 38 per cent year-on-year. The surge in exports was mainly driven by the agriculture-forestry-fisheries, and processing industries, which increased nearly 97 per cent and 38 per cent, respectively. The export of phones and components also rose by over 56 per cent to nearly $6 billion in January. Vit Nam's agricultural products maintained their strong export performance, with coffee prices surging by more than 35 per cent and rice prices escalating by 33.5 per cent year-on-year. The US remained Vit Nam's largest export market, with recorded imports of $9.6 billion, while traditional export markets such as China, the EU, and ASEAN also showed growth. Vit Nam's imports in January totaled more than $30.6 billion, with China remaining the largest exporter at nearly $11 billion. The trade balance continued to exhibit a surplus of $2.9 billion, with significant surpluses with the US and China. The Ministry of Industry and Trade anticipated challenges for this year's exports due to increased transport costs and escalating political tensions in the world. The ministry plans to focus on stimulating domestic consumption, promoting production, and monitoring market developments to avoid shortages or disruptions in the supply of goods. — VNS

Summary4

Vit Nam's exports reached nearly US$33.6 billion in January, 2024, a 42 per cent surge over the same period of 2023. Vit Nam's total imports and exports of goods in the first month surpassed $64 billion, up 38 per cent year-on-year. The surge in exports was mainly driven by the agriculture-forestry-fisheries, and processing industries, which increased nearly 97 per cent and 38 per cent, respectively. The export of phones and components also rose by over 56 per cent to nearly $6 billion. The US remained Vit Nam's largest export market, with recorded imports of $9.6 billion, while traditional export markets such as China, the EU, and ASEAN also showed growth. Vit Nam's imports in January totaled more than $30.6 billion, with China remaining the largest exporter at nearly $11 billion.

Text5

Oman records trade balance surplus of OMR 6.997 billion Muscat: The trade balance of the Sultanate of Oman recorded a surplus of OMR 6.997 billion at the end of November 2023, compared to a surplus of OMR 9.587 billion during the same period in 2022, according to preliminary statistics issued by the National Center for Statistics and Information. Statistics indicate that the value of merchandise exports at the end of last November reached OMR 20.636 billion, declining by 11.4 per cent compared to the same period in 2022, which recorded OMR 23.289 billion. The value of merchandise imports to the Sultanate of Oman amounted to OMR 13.639 billion, a decrease of 0.5 per cent at the end of November 2023 compared to the same period last year, which amounted to OMR 13.702 billion. The decline in the value of exports is mainly due to the decline in the value of the Sultanate of Oman's oil and gas exports, which now amounts to OMR 12.525 billion. At the end of November 2022, it amounted to OMR 15.215 billion. On the other hand, the value of the Sultanate of Oman's exports of crude oil amounted to OMR 8.9 billion by the end of November 2023, recording a decrease of 17.3 per cent compared to the same period of the previous year. The value of refined oil exports declined to OMR 1.3 billion, by 23.5 per cent. The value of the Sultanate's exports also decreased. Oman's LNG supply reached OMR 2.3 billion, or 15.7 per cent, compared to the end of November 2022, which amounted to OMR 2.7 billion. Statistics also revealed a decline in the value of non-oil merchandise exports by 1.4 percent at the end of November 2023, reaching OMR 6.767 billion, compared to the end of November 2022, when they recorded OMR 6.866 billion. Mineral products had the highest value among non-oil commodity exports, reaching 2.5 billion, an increase of 20.3 per cent, followed by ordinary metals and their products. The value of the Sultanate of Oman's exports of plastic and its products, rubber and its products also decreased to OMR 814 million, while exports of live animals and their products increased by 23 per cent to reach OMR 365 million, The value of exports of other products reached OMR 948 million. The Kingdom of Saudi Arabia topped the trade exchange transactions in non-oil exports, as their value at the end of November 2023 amounted to about OMR 981 million, an increase of 29.7 per cent from the end of November 2022, while the United Arab Emirates topped the trade exchange transactions in re-exports from the Sultanate of Oman, where the value of re-exports to it amounted to OMR 465 million at the end of last November. The United Arab Emirates also ranked first on the list of the countries exporting the most to the Sultanate of Oman, with a value of OMR 3.6 billion, a decrease of 8.6 per cent from the end of November 2022.

Summary5

The trade balance of the Sultanate of Oman recorded a surplus of OMR 6.997 bn at the end of November 2023. The value of merchandise exports at the end of November reached OMR 20.636 billion, declining by 11.4 per cent compared to the same period in 2022. The value of merchandise imports amounted to OMR 13.639 billion, a decrease of 0.5 per cent. The decline in the value of exports is mainly due to the decline in the value of oil and gas exports, which now amounts to OMR 12.525 billion. The value of the Oman's exports of crude oil amounted to OMR 8.9 billion by the end of November 2023, a decrease of 17.3 per cent. There was also a decline in the value of non-oil merchandise exports by 1.4 per cent, reaching OMR 6.767 billion. Mineral products had the highest value among non-oil commodity exports, reaching 2.5 billion, an increase of 20.3 per cent, followed by ordinary metals and their products. The value of the Sultanate of Oman's exports of plastic and its products, rubber and its products also decreased to OMR 814 million, while exports of live animals and their products increased by 23 per cent to reach OMR 365 million. The Kingdom of Saudi Arabia topped the trade exchange transactions in non-oil exports, as their value at the end of November 2023 amounted to about OMR 981 million, an increase of 29.7 per cent, while the UAE topped the trade exchange transactions in re-exports from the Sultanate of Oman, where the value of re-exports to it amounted to OMR 465 million at the end of November. The UAE also ranked first on the list of the countries exporting the most to the Sultanate of Oman, with a value of OMR 3.6 billion (-8.6% year-on-year).

B mt5 config

"_name_or_path": "csebuetnlp/mT5_multilingual_XLSum", "architectures": "MT5ForConditionalGeneration", "classifier_dropout": 0.0, "d_ff": 2048, "d_kv": 64, "d_model": 768, "decoder_start_token_id": 0, "dense_act_fn": "gelu_new", "dropout_rate": 0.1, "eos_token_id": 1, "feed_forward_proj": "gated-gelu", "initializer_factor": 1.0, "is_encoder_decoder": true, "is_gated_act": true, "layer_norm_epsilon": 1e-06, "length_penalty": 0.6, "max_length": 200, "model_type": "mt5", "no_repeat_ngram_size": 2, "num_beams": 4, "num_decoder_layers": 12, "num_heads": 12, "num_layers": 12, "output_past": true, "pad_token_id": 0, "relative_attention_max_distance": 128, "relative_attention_num_buckets": 32, "tie_word_embeddings": false, "tokenizer_class": "T5Tokenizer", "torch_dtype": "float32", "transformers_version": "4.34.1", "use_cache": true, "vocab_size": 250112

C Comparison of complex manual and model-generated summary

News link to iz.ru

Россия и Китай намерены продолжать углубление взаимовыгодного сотрудничества в различных форматах на фоне нестабильности мировой экономики и западных санкций. Об этом говорится в совместном заявлении обеих стран по итогам девятого китайско-российского финансового диалога в Пекине, которое было принято 18 декабря. «Стороны будут углублять сотрудничество в рамках многосторонних форматов, включая G20, БРИКС и АТЭС, с целью ограничения рисков, связанных с геополитической и геоэкономической фрагментацией, последовательно продвигать экономическую глобализацию и поддерживать стабильность и бесперебойность глобальных промышленных цепочек и цепочек поставок», — цитирует данное заявление «РИА Новости». Отмечается, что развитие мировой экономики пока что остается неопределенным в связи с продолжительной и высокой инфляцией, последствиями пандемии коронавируса, геополитической напряженностью и политически мотивированными ограничениями. В связи с этим стороны договорились придерживаться в рамках сотрудничества принципов взаимной выгоды, а также подтвердили приверженность усилению экономических связей с помощью координации макроэкономической политики. Кроме того, РФ и КНР выразили поддержку расширения БРИКС, которое названо отправной точкой для наращивания усилий по финансово-экономическому сотрудничеству стран объединения. Ранее в этот день вице-премьер РФ Дмитрий Чернышенко заявил после заседания российско-китайской межправкомиссии по подготовке регулярных встреч глав правительств, что Россия и Китай планируют к 2030 году увеличить объем товарооборота до $300 млрд. По его словам, за 11 месяцев 2023 года торговый оборот обеих стран уже превысил $201 млрд. До этого, 15 декабря, полпред президента РФ в Дальневосточном федеральном округе (ДФО) Юрий Трутнев сообщил, что за 10 лет товарооборот между дальневосточными регионами России и северо-восточными провинциями Китая увеличился вдвое. По его данным, в 2022 году данный показатель составил $22 млрд. Ожидается, что в 2023-м он достигнет $27 млрд. В тот же день вице-премьер РФ Александр Новак сообщил, что Россия и Китай договорились о сотрудничестве в области технологий по производству водорода. В свою очередь, премьер Госсовета КНР Дин Сюэсян после встречи с Новаком также отметил, что энергетическое сотрудничество двух стран в текущем году активно развивалось и достигло новых результатов даже в условиях сложной и напряженной международной обстановки. 7 декабря президент России Владимир Путин заявил на пленарном заседании форума «Россия зовет!», что Москва готова к всестороннему сотрудничеству с Пекином. Он подчеркнул, что РФ и КНР думают о перспективах и будущем мироустройстве, а также отходят от модели «купил-продал» в двусторонних отношениях

Manual summary: Россия и Китай подтвердили планы углублять взаимовыгодное сотрудничество. Россия и Китай намерены продолжать углубление взаимовыгодного сотрудничества в различных форматах на фоне нестабильности мировой экономики и западных санкций. Об этом говорится в совместном заявлении обеих стран по итогам девятого китайско-российского финансового диалога в Пекине, которое было принято 18 декабря. Ранее в этот день вице-премьер РФ Дмитрий Чернышенко заявил после заседания российско-китайской межправкомиссии по подготовке регулярных встреч глав правительств, что Россия и Китай планируют к 2030 году увеличить объем товарооборота до $300 млрд. До этого, 15 декабря, полпред президента РФ в Дальневосточном федеральном округе (ДФО) Юрий Трутнев сообщил, что за 10 лет товарооборот между дальневосточными регионами России и северо-восточными провинциями Китая увеличился вдвое. В тот же день вице-премьер РФ Александр Новак сообщил, что Россия и Китай договорились о сотрудничестве в области технологий по производству водорода.

Simple-model summary: Россия и Китай намерены продолжать углубление взаимовыгодного сотрудничества в различных форматах на фоне нестабильности мировой экономики. Об этом говорится в совместном заявлении обеих стран по итогам девятого китайско-российского финансового диалога в Пекине, которое было принято 18 декабря.

The best TinyLlama-model summary: Россия и Китай надеется продолжать усиленную выгодную инициативу и в течение намеченного года объединять во взаимовыгодные отношения по инициативе вице-премьер РФ Дмитрия Чернышенко. До этого с марта 2023 года обе страны обсуждают подготовку вновь созданных межправительственных консультативных комитетов по подготовке санкционных рамок и обсуждают договоренности об общих подходах к реализации генерующих взаимоволивостью направленных направленных межправительственных задач. Позже им также были подготовлены и приняты соответствующие документы законы о государственных контрактах по экономиче

The Mistral-model summary: Россия и Китай намерены углублять экономическое сотрудничество. Об этом говорится в совместном заявлении РФ и КНР по итогам девятого китайско-российского финансового диалога. Около $201 млрд Объём товарооборота между РФ и КНР по итогу 11 месяцев 2023 года составил $201 млрд, показатель за 2022 год составил $22 млрд. К 2030 году между сторонами планируется возможность увеличить объем товарооборота до $300 млрд. Двустороннее сотрудничество разви

Mt5: did not convey the necessary indicators correctly.

TinyLlama: provided excessive details on insignificant information.

Mistral: $22 billion is a metric specific to the Far Eastern region, not Russia.