International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Tabular Data to Text: A Transformative Approach

Kotagiri Hanvitha

Computer Science and Engineering Department, GMR Institute of Technology, Rajam, Andhra Pradesh, India

ABSTRACT –

This abstract introduces an approach that leverages transformer techniques to seamlessly convert tabular data into coherent and informative textual descriptions. Table-to-text generation has attracted widespread attention, aiming to help humans better understand tabular data. The system's core methodology involves processing input tables to find out underlying patterns, identify significant attributes, and extract relevant details. By automating the translation of structured data into human-readable narratives, the system enhances accessibility and comprehension, benefiting data analysts and decision-makers. Incorporating transformers in table-to-text generation has sparked a good shift, enabling seamless conversion of structured tabular data into coherent natural language text. In conclusion, this abstract presents an innovative solution that employs transformer techniques to revolutionize the conversion of tabular data into insightful textual narratives. By flawlessly integrating attention mechanism and natural language processing, this approach promises to enhance data communication, interpretation, and decision-making across a range of industries and applications.

Keywords - *Table to Text Generation, Natural Language Processing (NLP), Transformers, Attention Mechanism.*

1. Introduction

In today's world of using data to make things better, three important technologies are making a big impact. These are Transformers, Attention Mechanisms, and Natural Language Processing (NLP). This introduction explains how these three things are changing the way we understand and use complicated data. Transformers were first created to help with language tasks but have since become a crucial part of modern artificial intelligence (AI). They can understand complex relationships in data, not just in languages but in many different areas. When combined with Attention Mechanisms – inspired by how our brains focus on important things – Transformers become even more powerful. This is like how we humans can pick out important details from a lot of information.

Natural Language Processing has also gotten more advanced. It lets machines not only understand human language but also generate text that sounds very human-like. This helps bridge the gap between raw data and useful insights. When you bring NLP together with Transformers and Attention Mechanisms, AI systems can do more than just understand – they can also talk and create text that makes sense in different situations. The mix of Transformers, Attention Mechanisms, and NLP is changing the way we analyze data, communicate, and get value from big sets of information. As we explore further, we'll see how these technologies work together to give us new abilities and insights in our world of data. This trio is not just about understanding data; it's about making AI systems communicate and create context-aware text, which was not easy before. This trio is making a big difference, going beyond normal ways of processing data. It helps AI systems not only understand language and data relationships but also create meaningful insights. This shift is important for many industries, making predictions more accurate and changing how we interact with machines. As we move into this time of tech teamwork, the combination of Transformers, Attention Mechanisms, and NLP shows how AI keeps growing, bringing us to a time where dealing with complex data becomes a part of our everyday lives.

2. Literature Survey

Gatti et al. (2022) introduce VisToT, a project aiming to convert structured data into readable text through NLP for both tables and images. The approach addresses out-of-vocabulary challenges by directly copying words from the table, enhancing text quality through a sophisticated text deliberation method using transformers and attention mechanisms. The combination of NLP, attention mechanisms, and copying strategies results in a robust solution for transforming structured data into coherent and human-readable narratives. Chen et al. (2023) introduce TASD, a novel approach that combines large-scale pretrained language models with deliberation mechanisms for neural table-to-text generation. TASD utilizes a pointer generator model, achieving competitive performance on the numericNLG dataset with a BLEU score of 0.82 and a METEOR score of 0.77. The model outperforms baselines in accuracy at the word level, recall of the sequence, and fluency of sentences [1]. Kasner et al. (2023) present TABGENIE, a versatile toolkit for table-to-text generation. Leveraging Huggingface Datasets, TABGENIE provides debugging and data visualization tools, facilitating experimentation and development in the domain of table-to-text generation [2]. Parikh et al. (2020) introduce the TOTTO dataset, focusing on generating one-sentence descriptions from Wikipedia tables. While emphasizing dataset creation and evaluation metrics, the paper mentions the success of BERT-to-BERT models and pointer generator models without providing specific accuracy percentages [3]. Yang et al. (2021) address table-to-text challenges, focusing on accurate content copying. Their transformer-based model outperforms baselines on WIKIBIO and ROTOWIRE datasets, showcasing improvements in BLEU and ROUGE scores [4]. Ma et al. (2019) present a two-stage model for low-resource table-to-text generation, achieving a significantly higher BLEU score compared to the baseline. Though specific accuracy percentages are not provided, the paper suggests enhanced accuracy based on performance improvements [5]. Andrejczuk et al. (2022) introduce TABT5, achieving state-of-the-art results in various tasks, including spread sheet formula prediction, question-answering, and data-to-text generation [6] . Dhingra et al. (2019) propose the PAR-ENT metric for table-to-text precision evaluation. PG-Net stands out as a top performer, demonstrating high F-scores and accuracy in generating text from tables [7]. Hu et al. (2023) aim to enhance user-controlled table-to-text generation using fine-tuning with simulated noisy user inputs. The BART architecture-based models achieve improvements in BLEU points on both clean and noisy test cases [8]. Wang et al. (2021) present SANA, a two-stage approach for table-to-text generation, addressing limitations in traditional autoregressive methods. SANA achieves a remarkable PARENT-T recall on WikiPerson, indicating substantial improvement [9]. Li et al. (2023) propose Plan-then-Seam (PTS), a non-autoregressive table-to-text model, achieving significant speedup for inference time while maintaining comparable performance against strong two-stage competitors [10]. Gong et al. (2019) propose a hierarchical encoder addressing the inadequacy of one-dimensional table representation, outperforming strong baselines in BLEU score [11]. Perlitz et al. (2022) introduce DEVTC, a diversity-enhancing scheme for table-to-text generation, producing diverse valid outputs and demonstrating proficiency in LT-control [12]. Wu et al. (2022) address challenges in medical scientific table-to-text generation, proposing a two-step architecture with a Transformer encoder-decoder. The method outperforms baseline approaches, showcasing the effectiveness of pre-training on scientific medical data [13]. Liu et al. (2019) introduce AlignNet, addressing the task of table-to-text NLG with unseen schemas. The model outperforms baseline methods by a large margin [14]. Wang et al. (2020) propose a Transformer-based framework for faithful neural table-to-text generation with content-matching constraints. The proposed framework significantly outperforms existing models in terms of automatic evaluation scores [15].

3. Methodology

Dataset Description

The paper utilizes the WIKILANDMARKS dataset, specifically designed for Vision-augmented Table-To-Text Generation (VISTOT) in the tourism domain. This multimodal dataset encompasses 73,084 distinct world landmarks, incorporating tables, associated images, and descriptive sentences for each landmark. WIKILANDMARKS serves as a comprehensive resource, facilitating the exploration of the synergy between textual and visual information in generating contextually rich descriptions. With a focus on the tourism domain, the dataset enhances the understanding of how vision-augmented approaches can contribute to the generation of informative and engaging content. The inclusion of both structured tables and visual data in WIKILANDMARKS provides a holistic perspective, making it a valuable asset for advancing research in multimodal table-to-text generation tasks.

Preprocessing

In the data preprocessing phase of the VISTOT task, the authors begin by creating a novel multimodal dataset called WIKILAND-MARKS, comprising tables, images, and descriptive sentences for 73,084 world landmarks. They experiment with various methods, including Faster RCNN, ViT, CLIP-ViT, and Swin, for encoding both tables and images to accurately handle joint information. The proposed VT3 model is initialized from the pretrained BART language model but undergoes additional pre-training to strengthen the relationship between vision and language data and learn richer features. The authors introduce three pre-training strategies, including image-table matching, masked value modeling, and image captioning, to enhance the model's ability to generate natural language text conditioned on multimodal data. Overall, this phase aims to prepare the dataset and model for effective table-to-text generation.

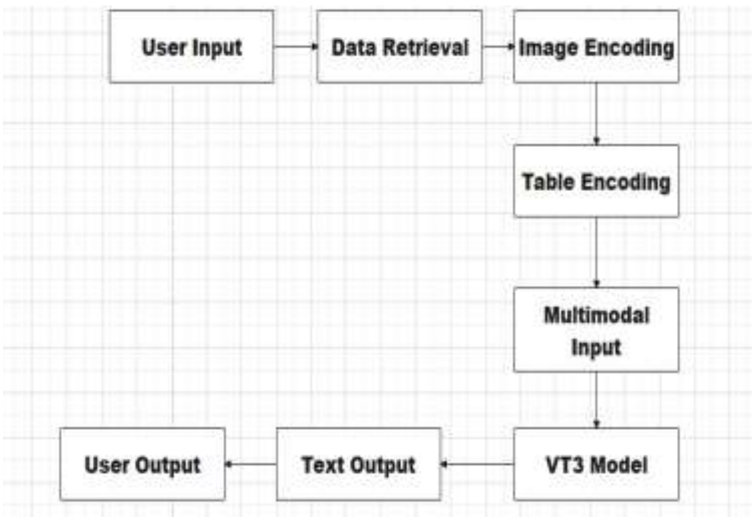

Fig. Structure

Architecture

The model introduces the Visual-Tabular Data-to-Text Transformer (VT3), a neural architecture based on the BART model, designed for multimodal data processing. Initialized from a pretrained BART language model, VT3 combines visual and tabular information in an encoder-decoder Transformer framework. Novel pretraining strategies, including image-table matching, masked value modeling, and image captioning, enhance its ability to generate precise sentences conditioned on multimodal inputs. The model's architecture captures essential features from text data, demonstrating its effectiveness in natural language generation. Extensive analyses and experiments underscore the significance of visual cues from images in handling incomplete or sparse tables. VT3 surpasses models relying solely on tabular data for table-to-text generation, highlighting the advantage of incorporating visual information. This research contributes to the advancement of table-to-text generation by leveraging the synergies between visual and tabular data, showcasing improved performance and addressing challenges associated with diverse data representations. The findings emphasize the potential of multimodal approaches for more accurate and contextually rich generation of human-readable narratives from structured data sources.

Evaluation metrics

Metrics including BLEU and METEOR were utilized to gauge the efficacy of the system.

BLEU (Bilingual Evaluation Understudy) evaluates the overlap of n-grams (contiguous sequences of n items, usually words) between the candidate translation and the reference translations.

$$\mathbf{BLEU=BP\times exp}\left(\sum_{n=1}^{N} w_n \cdot \log(p_n)\right)$$

where:

BP (Brevity Penalty): Corrects for short translations.

N: Maximum n-gram order considered (usually up to 4).

Wn: Weight for precision at n-grams.

Pn: Modified precision at n-grams, the ratio of matching n-grams in the candidate to the total n-grams in the candidate.

The **METEOR** (Metric for Evaluation of Translation with Explicit ORdering) assesses machine-generated translations using precision, recall, and a penalty term for unigram matching in its formula.

$$\mathbf{METEOR}=(1-\beta)\cdot \mathrm{precision}+\beta\cdot \mathrm{recall}-\gamma\cdot \mathrm{Penalty}$$

Where β (precision and recall importance) is commonly set to 1.0, γ (penalty for unigram matching) is often set to 0.5, and Penalty_count counts unigrams in the candidate with exact matches in the reference.

4. Results and Discussions

The performance summary of table-to-text generation models reveals diverse strengths. VT3, leveraging Transformers, Attention Mechanism, and LSTM, excels in generating descriptive sentences from multimodal data, achieving high BLEU and METEOR scores on WikiLandMarks. TASATG demonstrates moderate performance on TOTTO with its application of Multi-Head Attention and Cross Attention mechanisms. Pointer-Generator, employing Seq2Seq and Bert-based approaches, competes well on ToTTo. PIVOT, emphasizing Attention Mechanism and Bi-LSTM, excels in WikiBio. OpenNMT, utilizing Transformers and Attention Mechanism, performs competitively on WikiPerson. The choice among these models depends on specific dataset characteristics and task requirements in table-to-text generation.

TABLE: Comparison of different Table-to-Text Generation models performed on datasets.

Model	**BLEU**	**METEOR**
VT3	30.2	53.5
TASATG	14.9	11.65
openNMT	24.56	22.37
Pointer-Generator	19.2	25.19
PIVOT	27.34	35.1

5. Conclusion

In conclusion, the integration of transformers in table-to-text generation heralds a groundbreaking leap in data communication. Focused on pattern recognition and information extraction, this methodology enhances cognitive understanding for analysts and decision-makers. By seamlessly blending transformers, attention mechanisms, and natural language processing, the approach streamlines data analysis, automating raw data translation into insightful narratives. This transformative dimension not only amplifies accessibility but expedites decision-making, promising a positive evolution in data interpretation methodologies. The amalgamation of transformer techniques fosters collaboration, offering a potent means for industries to unlock the full cognitive potential of structured data, ushering in a new era of efficiency and comprehension.

6.References

1. Chen, M., Lu, X., Xu, T., Li, Y., Zhou, J., Dou, D., & Xiong, H. (2023). Towards table-to-text generation with pretrained language model: A table structure understanding and text deliberating approach. arXiv preprint arXiv:2301.02071.
2. Kasner, Z., Garanina, E., Plátek, O., & Dušek, O. (2023). TabGenie: A Toolkit for Table-to-Text Generation. arXiv preprint arXiv:2302.14169.
3. Parikh, A. P., Wang, X., Gehrmann, S., Faruqui, M., Dhingra, B., Yang, D., & Das, D. (2020). ToTTo: A controlled table-to-text generation dataset. arXiv preprint arXiv:2004.14373.
4. Yang, Y., Cao, J., Wen, Y., & Zhang, P. (2021). Table to text generation with accurate content copying. Scientific reports, 11(1), 22750.
5. Ma, S., Yang, P., Liu, T., Li, P., Zhou, J., & Sun, X. (2019). Key fact as pivot: A two-stage model for low resource table-to-text generation. arXiv preprint arXiv:1908.03067.
6. Andrejczuk, E., Eisenschlos, J. M., Piccinno, F., Krichene, S., & Altun, Y. (2022). Table-to-text generation and pre-training with tabt5. arXiv preprint arXiv:2210.09162.
7. Dhingra, B., Faruqui, M., Parikh, A., Chang, M. W., Das, D., & Cohen, W. W. (2019). Handling divergent reference texts when evaluating table-to-text generation. arXiv preprint arXiv:1906.01081.
8. Hu, H., Liu, Y., Yu, Z., & Perez-Beltrachini, L. (2023). Improving User Controlled Table-To-Text Generation Robustness. arXiv preprint arXiv:2302.09820.
9. Wang, P., Lin, J., Yang, A., Zhou, C., Zhang, Y., Zhou, J., & Yang, H. (2021). Sketch and refine: Towards faithful and informative table-to-text generation. arXiv preprint arXiv:2105.14778.
10. Li, L., Geng, R., Fang, C., Li, B., Ma, C., Li, B., & Li, Y. (2023). Plan-then-Seam: Towards Efficient Table-to-Text Generation. arXiv preprint arXiv:2302.05138
11. Gong, H., Feng, X., Qin, B., & Liu, T. (2019). Table-to-text generation with effective hierarchical encoder on three dimensions (row, column and time). arXiv preprint arXiv:1909.02304.
12. Perlitz, Y., Ein-Dor, L., Sheinwald, D., Slonim, N., & Shmueli-Scheuer, M. (2022). Diversity enhanced table-to-text generation via type control. arXiv preprint arXiv:2205.10938.
13. Wu, H. Y., Zhang, J., Ive, J., Li, T., Gupta, V., Chen, B., & Guo, Y. (2022, October). Medical Scientific Table-to-Text Generation with Synthetic Data under Data Sparsity Constraint. In NeurIPS 2022 Workshop on Synthetic Data for Empowering ML Research.
14. Liu, T., Wei, W., & Wang, W. Y. (2019). Table-to-Text Natural Language Generation with Unseen Schemas. arXiv preprint arXiv:1911.03601.
15. Wang, Z., Wang, X., An, B., Yu, D., & Chen, C. (2020). Towards faithful neural table-to-text generation with content-matching constraints. arXiv preprint arXiv:2005.00969