EROS: Entity-Driven Controlled Policy Document Summarization

Joykirat Singh, Sehban Fazili, Rohan Jain, and Md Shad Akhtar

LREC-COLING 2024

Motivation

- Lengthy and Complex Policies: Privacy policies are often long and filled with technical and legal jargon, making them challenging for users to read and comprehend.
- User Empowerment: Privacy policy summarization aims to empower users by providing them with accessible information about their privacy rights and how their personal data is handled.
- Time Efficiency: Privacy policies can be time-consuming to read and understand in their entirety.
- Accessibility for Non-Experts: Privacy policies often contain technical terms and legal language that may be difficult for non-experts to comprehend.

What is a Policy Document?

- Formal statement outlining **rules, guidelines, or principles.**
- **Importance:** Promote transparency, accountability, and compliance.
- Most of us accept to the terms and conditions or policies **without reading them.**
    - **Reasons:**

I don't have time

I don't care!

Sample Policy

- When you register with the Site or use any of our Services, you may be asked to provide us with Personal Information. It is entirely optional to provide this information. If you do not provide the requested information, you may be unable to use some or all of the Site's features. To improve the Site and Services, we use cookies (a small text file placed on your computer to identify your computer and browser) and Web beacons (electronic files placed on a Web site that monitor usage). If you delete or disable cookies, some features of the Site or Services may not function properly. You have the option of making some of your Personal Information available to others. Users of the Site and commercial search engines such as Google, Yahoo!, and Bing may access the information you choose to make available. Some third-party services may provide us with information from your accounts, allowing us to improve and personalise your experience on the Site. When you visit our Site, we may allow third-party companies to serve ads and/or collect certain anonymous information …

What data are being captured?

Who is capturing?

Sample Policy

- When **you register with the Site or use any of our Services**, you may be asked to provide **us** with **Personal Information**. It is entirely optional to provide this information. If you do not provide the requested information, you may be unable to use some or all of the Site's features. To **improve the Site and Services**, we use **cookies** (a small text file placed on your computer to identify your computer and browser) and **Web beacons** (electronic files placed on a Web site that monitor usage). If you delete or disable cookies, some features of the Site or Services may not function properly. You have the option of making some of your **Personal Information** available to others. **Users of the Site** and commercial search engines such as **Google**, **Yahoo!**, and **Bing** may access the information you choose to make available. Some **third-party services** may provide **us** with information from your accounts, allowing us to **improve and personalise your experience on the Site**. When **you visit our Site**, we may allow **third-party companies** to serve ads and/or collect certain **anonymous information**
- **Data compulsory**: [e-mail, personally identifiable information];
  Data others: [personal information];
  Source Direct: [you, info@dayfinder.com];
  Target Direct: [we, us];
  Target Indirect: [service providers, third parties];
  Medium: [register with the Site or use any of our Services, cookies, web beacons, visit our site];
  Reason: [improve the site and services, improve and personalise your experience on the Site, Usage Data regarding the Site and Services].

Contribution

- The main contributions are as follows:
  - We propose a BART-based entity-driven controlled policy document summarization (EROS) to mitigate the concerns of general public over the data privacy and security issues.
  - To identify privacy-related relevant information in a policy document, we developed an entity extraction model, Entity Extraction from Policy Documents (EEPD).
  - We introduce a personalized loss function and a reinforcement learning framework using Proximal Policy Optimization (PPO) to manage the relevance and length of the generated summaries.
  - We introduce a new dataset (PD-Sum) of privacy policy documents with their summaries and privacy-bounded entities and rationales.

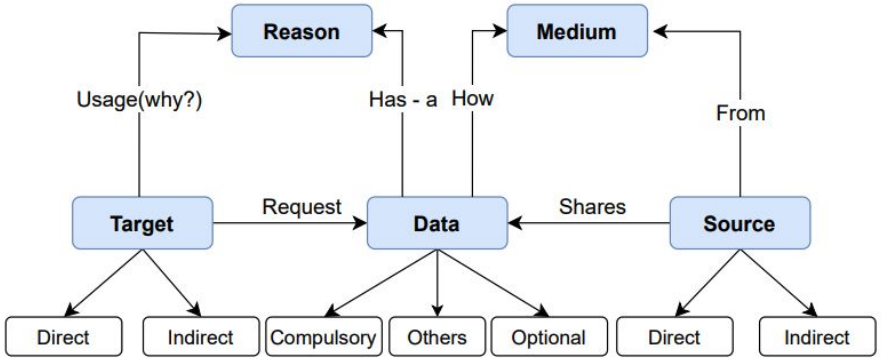
Dataset (PD-Sum)

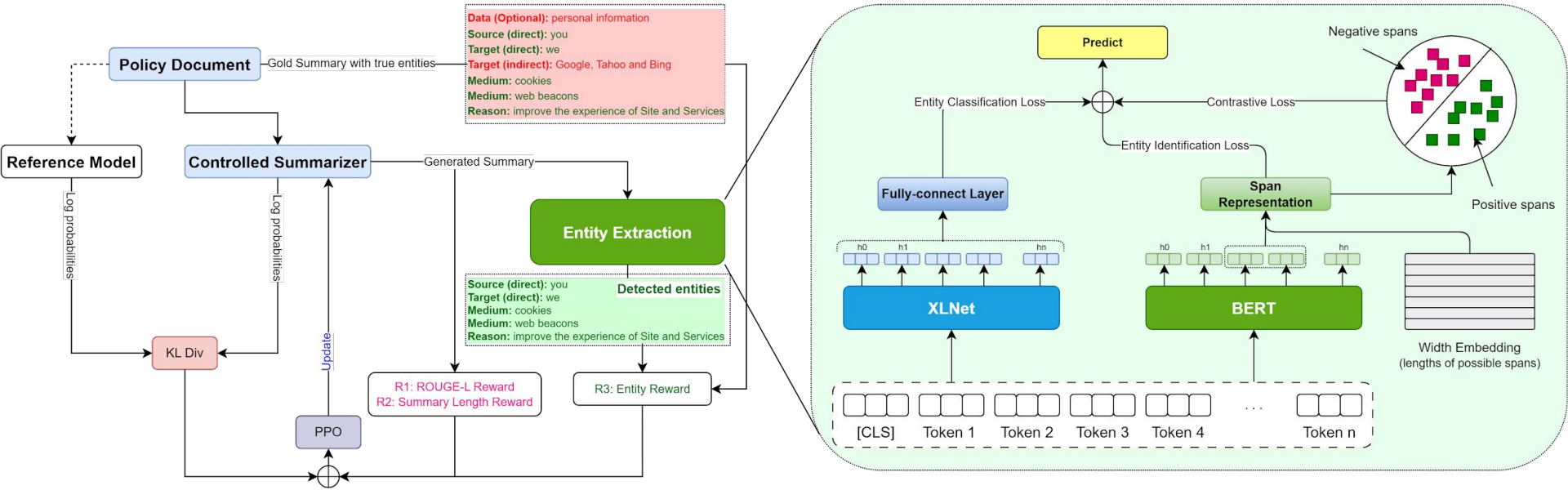
- Manually Annotated dataset PD-Sum
  - Manually annotated documents with adstractive summaries along with privacy related entities and their rationales. (0.74 cohen kappa Score)
  - 1921 policy documents.

Total Documents	*1920*
Total Paragraphs	*1,20,991*
avg_tokens_per_doc	*1707.3*
avg_tokens_per_summary	*228.24*

Dataset (PD-Sum)

- Data Compulsory : Data which is compulsory for the source to enter
- Data Optional : Data in which the source has the option to provide to the target
- Data Others : Information that belongs to the source
- Source Direct : The entity that directly provides the data to the target
- Source Indirect : The entity that indirectly provides the data to the target
- Target Direct : The entity to which the source directly provides the data
- Target Indirect : The entity to which the source indirectly provides the data
- Reason : Reason for why the data is being collected by the target
- Medium : How the data will be collected by the target





Training framework of EROS

Methodology - EROS

- Modified Loss function (cross entropy loss with a penalty component) :
  - Augmenting the traditional cross-entropy loss with a penalty component derived from the extracted entities.
  - Enables the BART model to comprehend the presence and importance of entities in the summaries.

$$\mathsf{CE} = -\sum(y \cdot \log(x))$$

$$TP = \sum_{e_i}(1.0 - \mathsf{step}(e_i \in S_G))$$

$$\mathcal{L}^s = \lambda \cdot \mathsf{CE} + (1 - \lambda) \cdot \mathsf{TP}$$

Gold entities (e_i)

Methodology - EROS

- Reward Calculation:
  - Coverage Reward (R1) : Ensures the readability of the generated summary.

$$R1 = \text{ROUGE-L}(S_G, S_R))$$

  - Conciseness Reward (R2) : Limits the model to generate an adequate length summary and avoid generating lengthy jargon.

$$R_2 = \frac{1 - |(len(S_G) - len(S_R))|}{max(len(S_G), len(S_R))}$$

  - Entity Reward (R3)

$$R_3 = \frac{E_{correct} - \beta * E_{incorrect}}{E_{total}}$$

Methodology - Entity Extraction

- Sentence - “we will collect name”, maximum span length - 4
- Possible spans : **{x_1, x_1}**, {x_1, x_2}, {x_1, x_3}, {x_1, x_4}, {x_2, x_2}, {x_2, x_3}, {x_2, x_4}, {x_3, x_3}, {x_3, x_4}, **{x_4, x_4}**

  {x_1, x_1} - Source Target, **{x_4, x_4}** - Data Compulsory

- XLNET - Increase in number of false positive without a binary classifier.

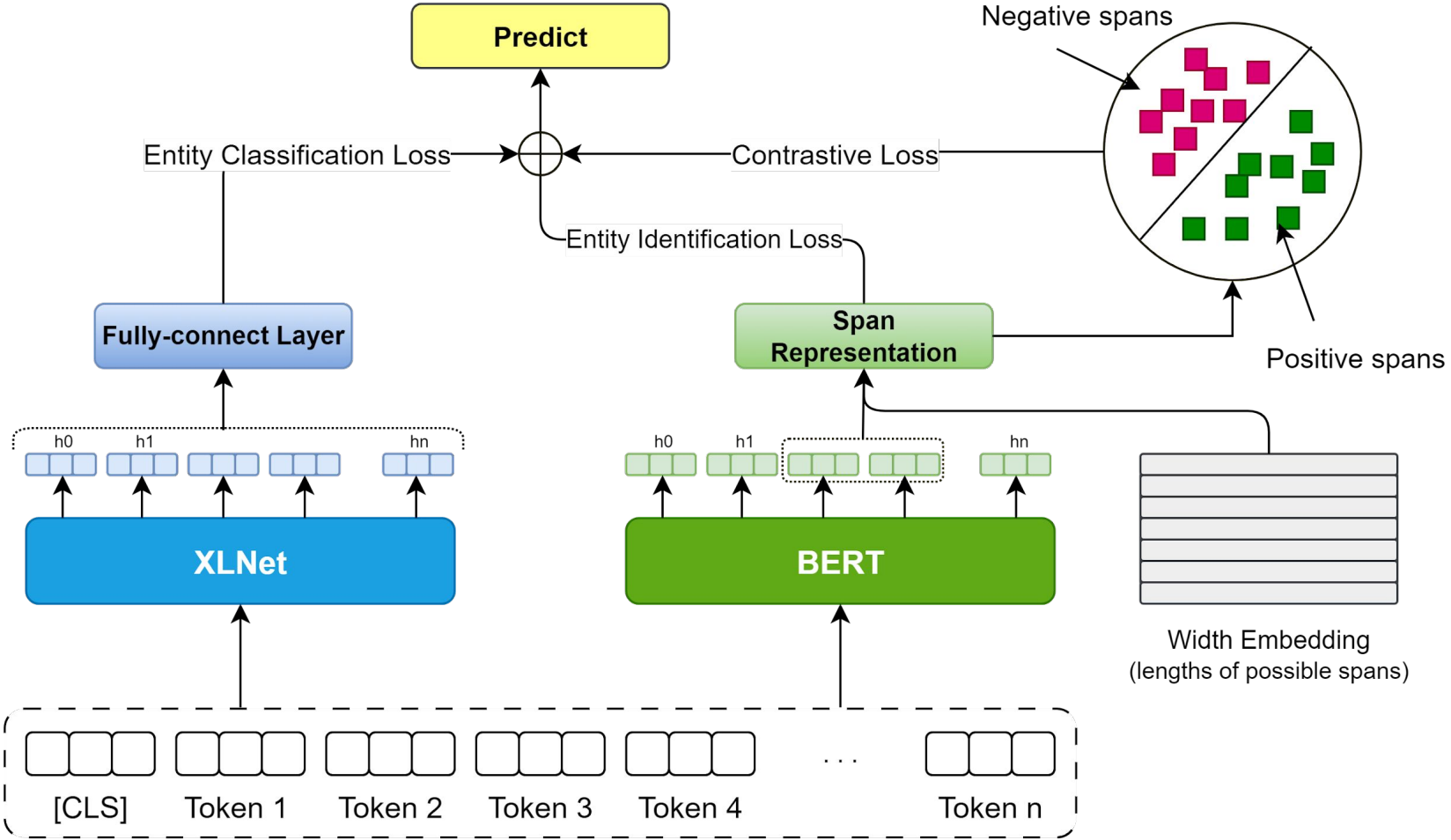
Methodology - Entity Extraction

- Contrastive Loss:
  - Compute Similarity score between pairs of input spans and minimise distance between similars pairs and maximise distance between dissimilar pairs.

$$\mathrm{P}\left(\mathbf{y} \mid s_i\right)=\frac{\operatorname{score}\left(\mathbf{s}_i, \mathbf{y}\right)}{\sum_{\mathbf{y}^{\prime} \in \mathcal{Y}} \operatorname{score}\left(\mathbf{s}_i, \mathbf{y}^{\prime}\right)}$$

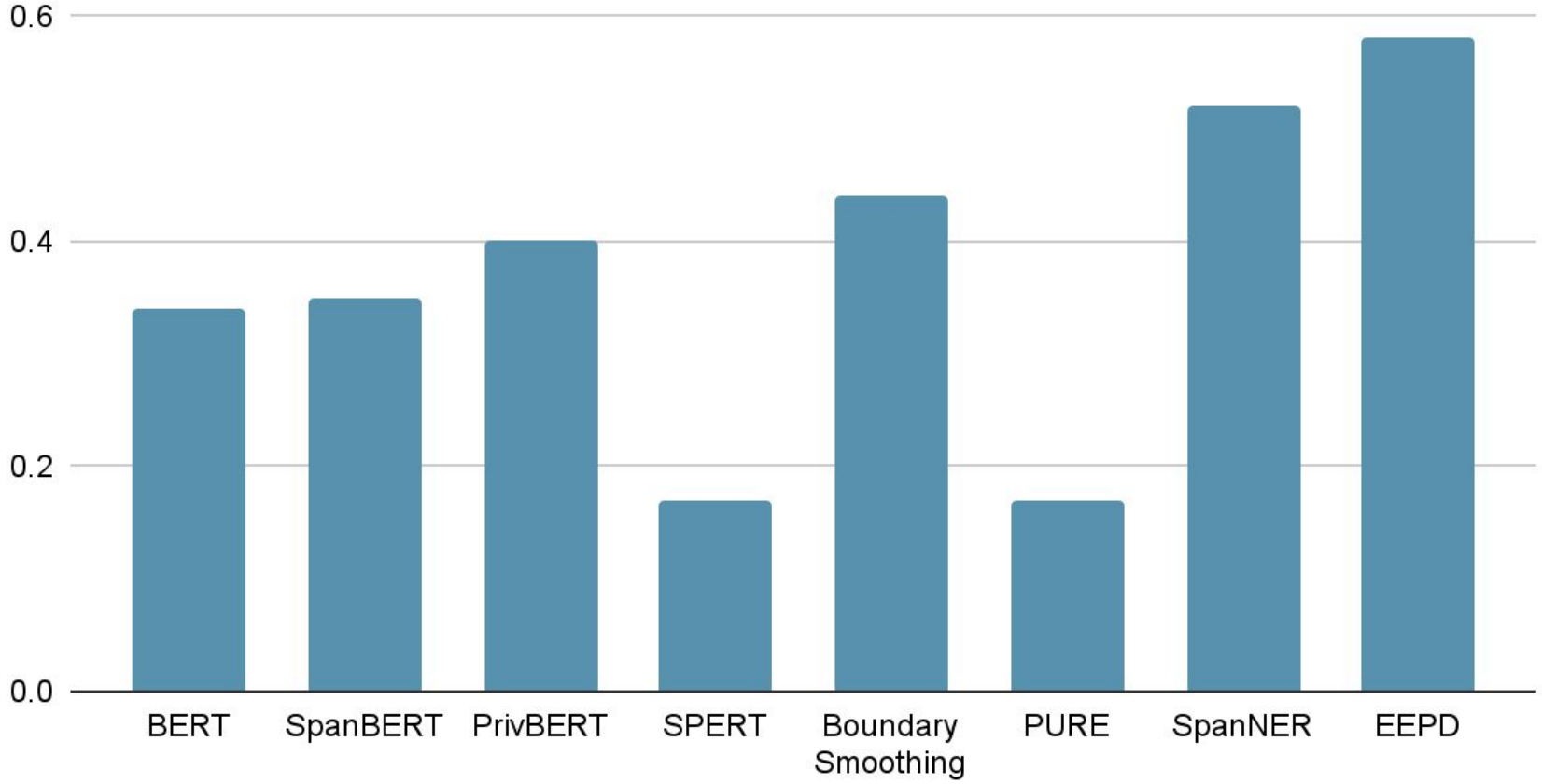
- Span Prediction:
  - Span representation are fed into softmax function, with three losses:
    - Entity Identification cross entropy loss (L_1)
    - Binary Classification Loss (L_2)
    - Contrastive Loss (L_3)

$$\mathcal{L}^e = \alpha_1 \ell_1 + \alpha_2 \ell_2 + \alpha_3 \ell_3, \quad \text{where} \sum \alpha_i = 1$$

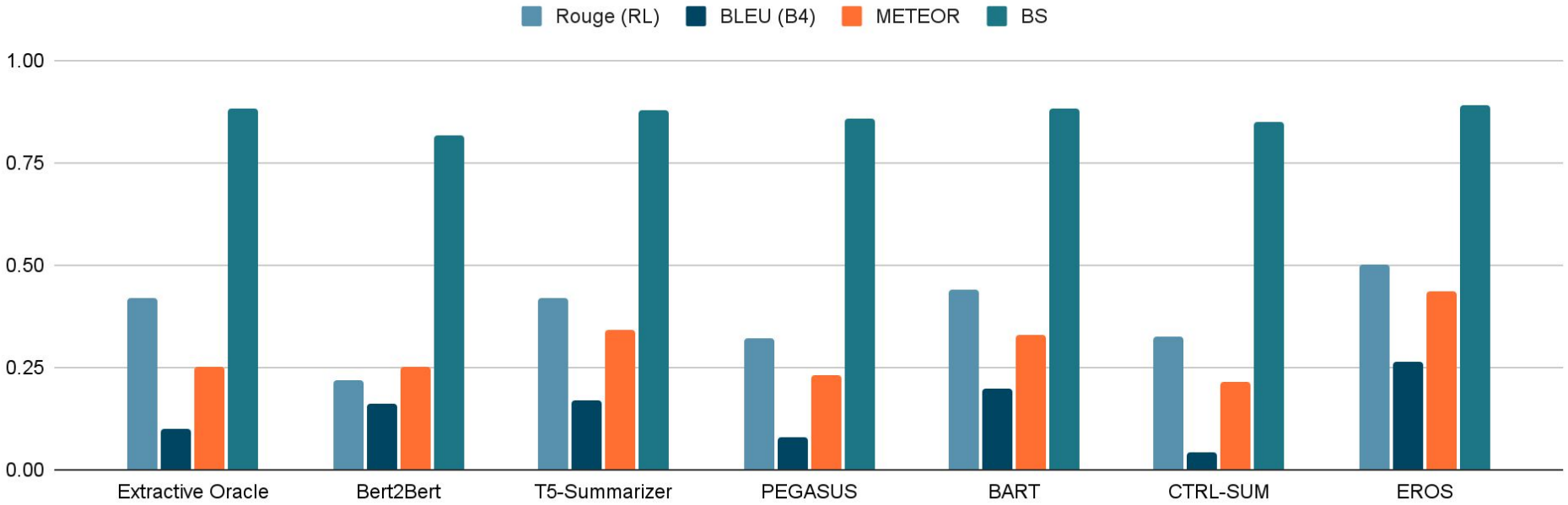


Model Overview of Entity Extraction Module

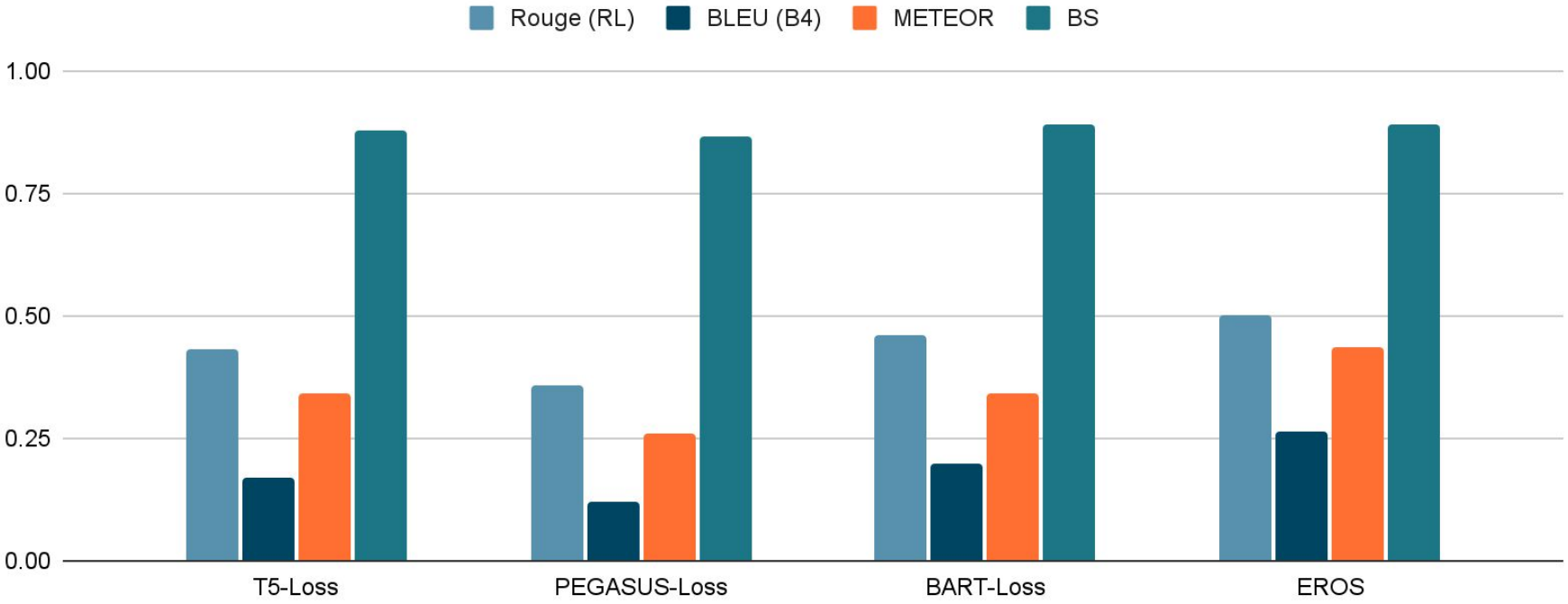
Results of Entity Extraction Module (F1 Score)



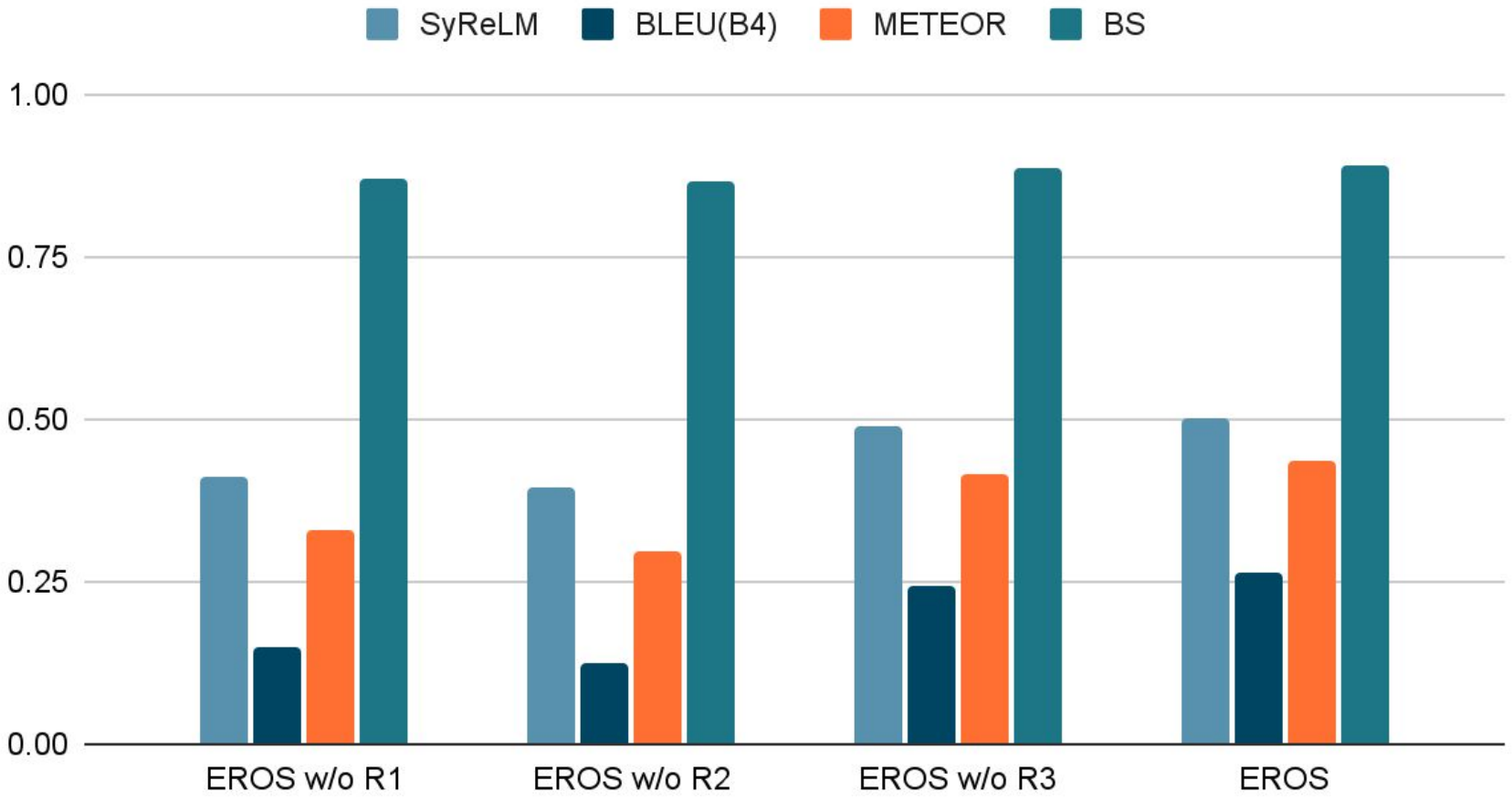
Result of EROS vs. Base models



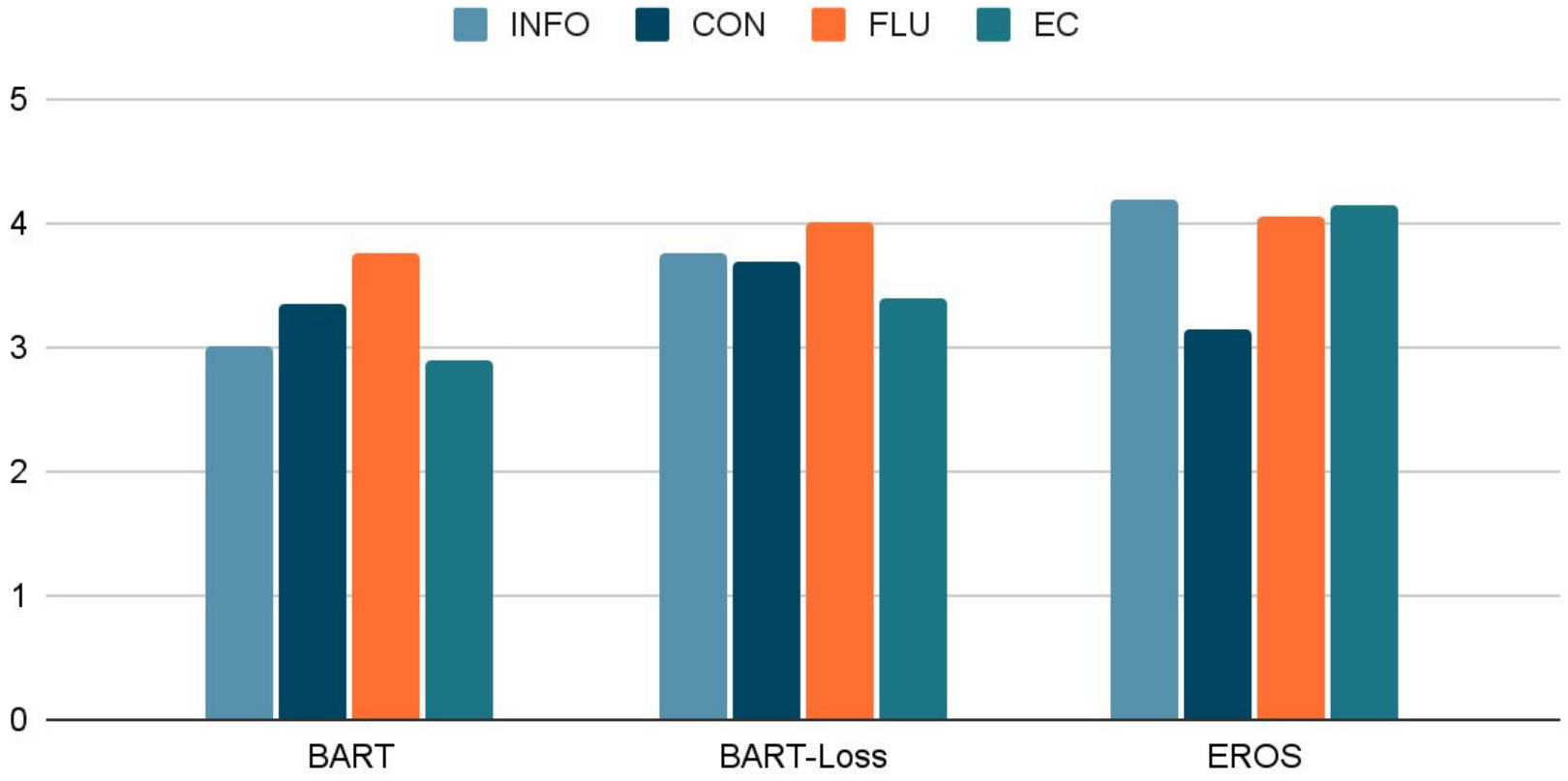
Result of EROS vs. Base models with modified loss



Ablation Study of EROS



Human evaluation on informativeness, conciseness, fluency, and entity coverage.



Key Takeaways

Novel approach to address challenge of generating controlled and information summaries.

Captured the essence of complex privacy policies.

Introduced a new Dataset.

State of the art performance on custom dataset.

Human Evaluation -> Generated summary from EROS surpassed in informativeness, fluency, grammatical correctness and inclusion of relevant entities.

Thanks for your attention
Questions?

joykiratsingh18@gmail.com

Code