

Analysis of Similarity Measures between Short Text for the NTCIR-12 Short Text Conversation Task

Kozo Chikai and Yuki Arase

Graduate school of information science and technology, Osaka University

Introduction

Goal

Given an input post, search appropriate replies from the pool.

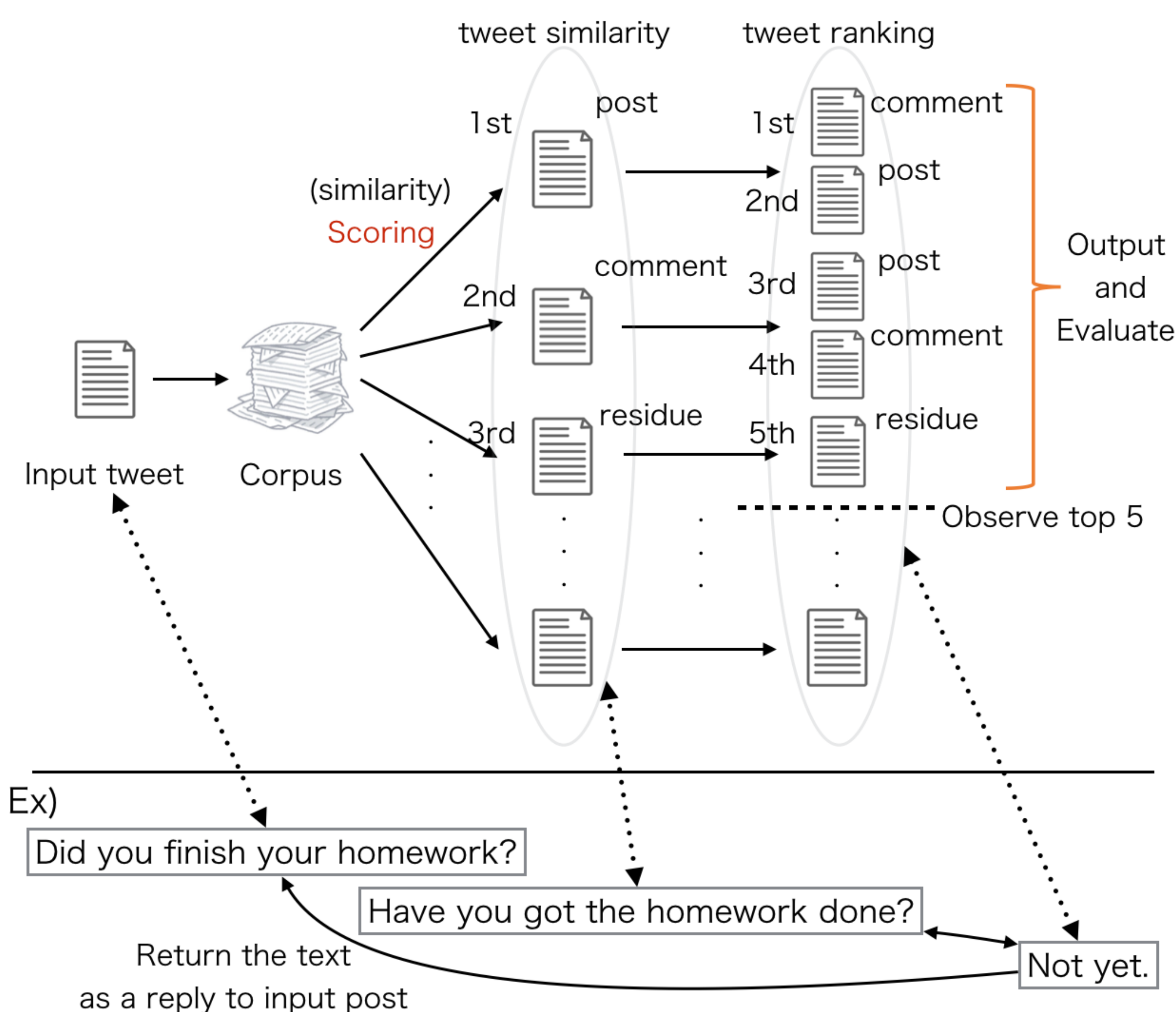
Approach

we compare the state-of-the-art methods for estimating text similarities to investigate their performance in handling short text, specially, under the scenario of short text conversation.

System Design

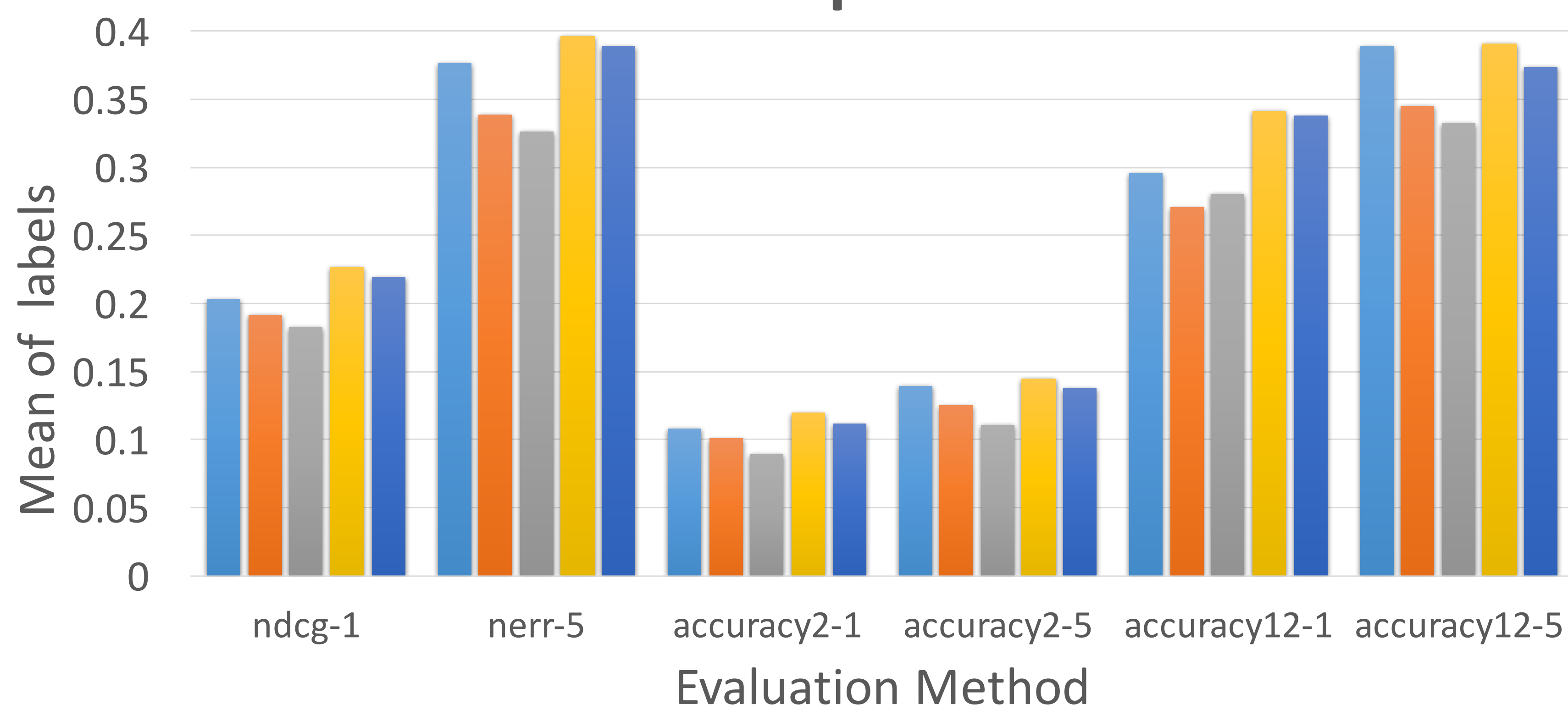
We implemented these five methods as Similarity Calculation.

- ① TF-IDF
- ② Weighted Text Matrix Factorization
- ③ Word2vec → TF-IDF
- ④ Random Forest → TF-IDF
- ⑤ Random Forest + TF-IDF



Results

Result of experiments



Mean of labels for each rank

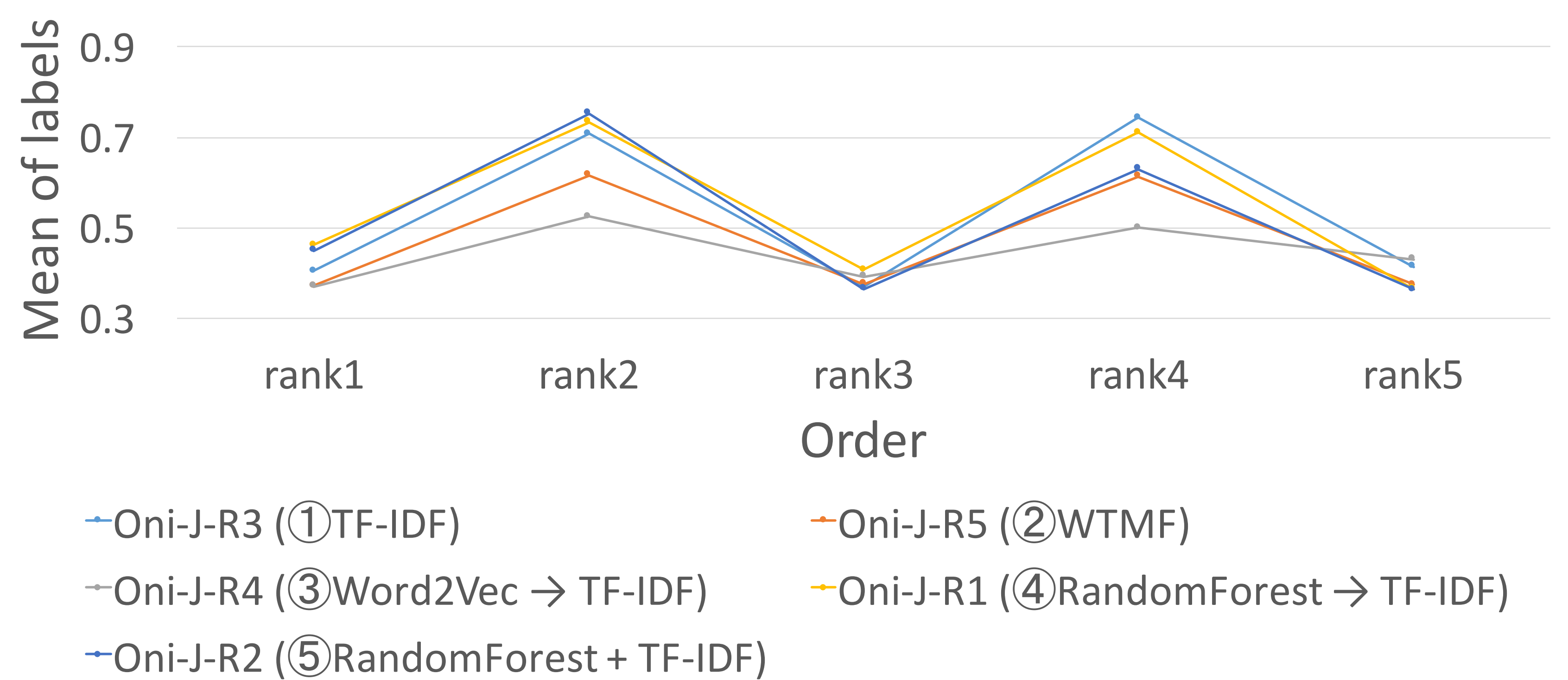


Table 1: Example of replies generated by Oni-J-R1

Input post	フォローありがとうございます！ Thank you for following!
rank1	こちらこそ、懐かしい画像をありがとうございました♪ よろしく願います！ Thank you too for sharing the nostalgic photos! Keep in touch.
rank2	RT ありがとです♪ フォローさせていただきました！ Thank you for retweeting. I had followed you!
rank3	お休みなさい♪ 今年も宜しく願います！ Good night. Wish we'll have fun together!
rank4	寝ます。おやすみなさい。 さっそく今日から仕事だ！ Going to bed. Good night. Gotta work from today!
rank5	んう いくう hmmm...

Table 2: Example of output from Oni-J-R3 and Oni-J-R4

Input post	LINE 使えんので用のある人はこっちでお願いします！ Cannot use LINE now. Please contact me via Twitter!	
Method	R3	R4
rank1	新しいアカウント作ったから教えてー！ Let me know your account! I've just created one.	PC で扱うようなでかい解像度で色々したりするんでなければいいよ十分な感じなんです？ It should be satisfactory unless you process high-resolution images that require PC-like performance.
rank2	LINE 使えるようになった？ Has LINE recovered?	んー、パソコンで使えた機能の大半はもう使えてるよね。単純に使える、ただけけど。 次の8コアのスマートフォン向けの cpu が出るからそれで単純に使える機能から卒業して、もうスマートフォンで十分になる機能は出るだろう Most functions in PCs are already covered by smart phones, although they are just functioning... They should be actually usable with the next 8-core CPU for smart phones.
rank3	パソコンの LINE ツンデレすぎるよ！ もう！ 早く携帯よとどけ！ LINE app for PC is hard to use... Get me a cell phone!	単発でやる環境に長くいたせいで、最近キャンペ参加できるようになったのですがなかなか上手い感じにキャンペ用キャラが組めないんですよ…… I cannot figure out how to organize my party for online campaign as I've been playing alone...