

**LAPORAN PRAKTIKUM  
SYSTEM INFORMATION RETRIEVAL  
RELEVANCE FEEDBACK**



**U N I V E R S I T A S  
AMIKOM PURWOKERTO**

23SP11C031  
Misno Sugianto

**PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM  
PURWOKERTO  
2023**

**Tugas Praktikum**

1. Modifikasi kode program! Apabila ditemukan Score = 0.0 dari hasil kalkulasi Relevance Feedback, maka nomor index dengan score 0.0 tersebut tidak perlu ditampilkan kepada pengguna.
2. Jelaskan minimal 3 alasan mengapa teknik relevance feedback jarang digunakan pada mesin pencarian berbasis web? Sertakan referensi pendukung anda!
3. Teknik relevance feedback dapat diimplementasikan pada model temu balik informasi apa saja, selain vector space model? Sertakan referensi pendukung anda!

## **Pembahasan**

1. Memodifikasi Kode program agar indeks dengan score 0. tidak ditampilkan. Kode awalnya seperti dibawah ini.

```
1  if __name__ == "__main__":
2      # Pengguna menginputkan query
3      query = input("Masukkan query: ")
4      # Menampilkan hasil awal Vector Space Model (VSM)
5      top_documents = vsm.calculate_vsm(query, corpus)
6
7      print("\n7 Peringkat Dokumen Teratas:")
8      for doc_idx, i in top_documents:
9          #if(i != 0.):
10             print(f"No. index {doc_idx}: Score {i}. {data[doc_idx]}")
11             print()
```

kode diatas akan menghasilkan result query seperti berikut apabila dijalankan

```
Masukkan query: surga
7 Peringkat Dokumen Teratas:
No. index 31: Score [0.44018973]. (32, 2, 25, 'Dan sampaikanlah berita gembira kepada mereka yang beriman dan berbuat baik, bahwa bagi mereka disediakan surga-surga yang mengalir sungai-sungai di dalamnya. Setiap mereka diberi rezeki buah-buahan dalam surga-surga itu, mereka mengatakan: "Inilah yang pernah diberikan kepada kami dahulu". Mereka diberi buah-buahan yang serupa dan untuk mereka di dalamnya ada isteri-isteri yang suci dan mereka kekal di dalamnya.')
```

```
No. index 88: Score [0.34549935]. (89, 2, 82, 'Dan orang-orang yang beriman serta beramal saleh, mereka itu penghuni surga; mereka kekal di dalamnya.')
```

```
No. index 44: Score [0.21062798]. (45, 2, 38, 'Kami berfirman: "Turunlah kamu semuanya dari surga itu! Kemudian jika datang petunjuk-Ku kepadamu, maka barang siapa yang mengikuti petunjuk-Ku, niscaya tidak ada kekhawatiran atas mereka, dan tidak (pula) mereka bersedih hati".')
```

```
No. index 42: Score [0.19981536]. (43, 2, 36, 'Lalu keduanya digelincirkan oleh syaitan dari surga itu dan dikeluarkan dari keadaan semula dan Kami berfirman: "Turunlah kamu! sebagian kamu menjadi musuh bagi yang lain, dan bagi kamu ada tempat kediaman di bumi, dan kesenangan hidup sampai waktu yang ditentukan.')
```

```
No. index 41: Score [0.1814211]. (42, 2, 35, 'Dan Kami berfirman: "Hai Adam, diamilah oleh kamu dan isterimu surga ini, dan makanlah makanan-makanannya yang banyak lagi baik dimana saja yang kamu sukai, dan janganlah kamu dekati pohon ini, yang menyebabkan kamu termasuk orang-orang yang zalim.')
```

```
No. index 0: Score [0.]. (1, 1, 1, 'Dengan menyebut nama Allah Yang Maha Pemurah lagi Maha Penyayang.')
```

```
No. index 1: Score [0.]. (2, 1, 2, 'Segala puji bagi Allah, Tuhan semesta alam.')
```

```
Masukkan nomor index dokumen relevan (pisahkan dengan spasi): |
```

kode diatas masih menghasilkan query dimana score dibawah 0.1 masih ditampilkan, sementara kita ingin indeks dengan score dibawahnya jangan ditampilkan. Perubahan kode yang dapat kita lakukan adalah dengan cara berikut :

```
1  print("\n7 Peringkat Dokumen Teratas:")
2  for doc_idx, i in top_documents:
3      if(i != 0. and i != 0.1):
4          print(f"No. index {doc_idx}: Score {i}. {data[doc_idx]}")
5          print()
```

Dengan kode diatas kita mencari dimana jika I bernilai tidak seperti 0. dan pembulatannya tidak sama dengan 0.1 maka tampilkan hasil dari I tersebut. Yang kemudian result dari kode diatas sebagai berikut :

```
Masukkan query: surga

7 Peringkat Dokumen Teratas:
No. index 31: Score [0.44018973]. (32, 2, 25, 'Dan sampaikanlah berita gembira kepada mereka yang beriman dan berbuat baik, bahwa bagi mereka disediakan surga-surga yang mengalir sungai-sungai di dalamnya. Setiap mereka diberi rezeki buah-buahan dalam surga-surga itu, mereka mengatakan: "Inilah yang pernah diberikan kepada kami dahulu". Mereka diberi buah-buahan yang serupa dan untuk mereka di dalamnya ada isteri-isteri yang suci dan mereka kekal di dalamnya.')
```

No. index 88: Score [0.34549935]. (89, 2, 82, 'Dan orang-orang yang beriman serta beramal saleh, mereka itu penghuni surga; mereka kekal di dalamnya.')

No. index 44: Score [0.21062798]. (45, 2, 38, 'Kami berfirman: "Turunlah kamu semuanya dari surga itu! Kemudian jika datang petunjuk-Ku kepadamu, maka barang siapa yang mengikuti petunjuk-Ku, niscaya tidak ada kekhawatiran atas mereka, dan tidak (pula) mereka bersedih hati".')

No. index 42: Score [0.19981536]. (43, 2, 36, 'Lalu keduanya digelincirkan oleh syaitan dari surga itu dan dikeluarkan dari keadaan semula dan Kami berfirman: "Turunlah kamu! sebagian kamu menjadi musuh bagi yang lain, dan bagi kamu ada tempat kediaman di bumi, dan kesenangan hidup sampai waktu yang ditentukan".')

No. index 41: Score [0.1814211]. (42, 2, 35, 'Dan Kami berfirman: "Hai Adam, diamilah oleh kamu dan isterimu surga ini, dan makanlah makanan-makanannya yang banyak lagi baik dimana saja yang kamu sukai, dan janganlah kamu dekati pohon ini, yang menyebabkan kamu termasuk orang-orang yang zalim.')

Masukkan nomor index dokumen relevan (pisahkan dengan spasi): |

Di hasil yang kita peroleh, indek yang memiliki score 0. dan pembulatannya 0.1 dihilangkan dan tidak ditampilkan.

2. Teknik umpan balik relevansi jarang digunakan di mesin pencari berbasis web karena beberapa alasan. Pertama, mengharuskan pengguna untuk memberikan umpan balik substansial pada daftar hasil tidak praktis dalam skenario pencarian web yang khas [1]. Kedua, pertanyaan pendek dan kabur, volume dokumen yang terus meningkat, dan pengguna yang tidak mengetahui kebutuhan informasi pasti mereka berkontribusi pada penurunan kinerja sistem IR, membuat umpan balik relevansi kurang efektif [2]. Terakhir, sebagian besar algoritma umpan balik relevansi menderita masalah kumpulan data yang tidak seimbang, di mana jumlah gambar yang tidak relevan jauh lebih besar daripada jumlah gambar yang relevan, yang mengarah ke hasil pengambilan yang terdegradasi [3]. Tantangan ini membuat sulit untuk menerapkan dan memanfaatkan teknik umpan balik relevansi di mesin pencari berbasis web.
3. Teknik umpan balik relevansi dapat diimplementasikan pada berbagai model pengambilan informasi, bukan hanya model ruang vektor. Misalnya, satu makalah mengusulkan pendekatan kNN yang mengintegrasikan umpan balik relevansi langsung ke dalam model peringkat ulang saraf, meningkatkan kinerjanya [4]. Makalah lain memperkenalkan informasi relevansi berbasis topik ke dalam model relevansi RM3 dan model Rocchio, membangun model yang umumnya lebih unggul dari model umpan balik relevansi semu [5]. Selain itu, makalah lain mengembangkan metode untuk umpan balik relevansi berdasarkan ruang kueri semantik, menggunakan metrik Riemann yang berasal dari distribusi probabilitas sampel positif [6]. Akhirnya, sebuah makalah mengusulkan memasukkan informasi semantik ke dalam umpan balik relevansi semu tradisional menggunakan model BERT, yang meningkatkan keakuratan sistem pengambilan [7].

## Referensi

- [1] K. Bi, Q. Ai, and W. B. Croft, "Iterative Relevance Feedback for Answer Passage Retrieval with Passage-level Semantic Match." arXiv, Dec. 20, 2018. Accessed: Dec. 13, 2023. [Online]. Available: <http://arxiv.org/abs/1812.08870>
- [2] M. Alam and K. Sadaf, "Relevance feedback versus web search document clustering," in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, Mar. 2015, pp. 1665–1669. Accessed: Dec. 13, 2023. [Online]. Available: <https://ieeexplore.ieee.org/document/7100530>
- [3] N. Bhosle and M. Kokare, "Random forest based long-term learning for content based image retrieval," in *2016 International Conference on Signal and Information Processing (IConSIP)*, Maharashtra State, India: IEEE, Oct. 2016, pp. 1–4. doi: 10.1109/ICONSIP.2016.7857468.
- [4] T. Baumgärtner, L. F. R. Ribeiro, N. Reimers, and I. Gurevych, "Incorporating Relevance Feedback for Information-Seeking Retrieval using Few-Shot Document Re-Ranking." arXiv, Oct. 19, 2022. Accessed: Dec. 13, 2023. [Online]. Available: <http://arxiv.org/abs/2210.10695>
- [5] "Pseudo-Relevance Feedback Method Based on the Topic Relevance Model," *Math. Probl. Eng.*, vol. 2022, pp. 1–6, Jul. 2022, doi: 10.1155/2022/1697950.
- [6] S. Santini, "Relevance Feedback with Latent Variables in Riemann spaces." arXiv, Jun. 15, 2019. Accessed: Dec. 13, 2023. [Online]. Available: <http://arxiv.org/abs/1906.06526>
- [7] "A Multi-Dimensional Semantic Pseudo-Relevance Feedback Information Retrieval Model," pp. 866–872, Nov. 2022, doi: 10.1109/WI-IAT55865.2022.00141.