Chapter 6
Discussion

This chapter discusses the results presented in chapter 5 obtained from the implementation of different frameworks proposed in chapter 4. The comparative analysis of the Present Search Engine (PSE) with the proposed and implemented frameworks is also discussed. Section 6.1 discusses the performance of Noise Removal Framework from the aspect of fixing the Domain. The performance of NRSIP framework and its comparison with PSE (Yahoo!) is discussed in section 6.2 and 6.3. Comparative analysis and performance evaluation of NRSIP-G framework (Generalized NRSIP Framework) and PSE (Yahoo!) is discussed in section 6.4. The retrieval effectiveness of Secured CAbMsISS and its comparative evaluation with PSE (Google) is illustrated in Section 6.5.

6.1 Noise Removal Framework (116)

Noise Removal Framework has been successfully tested from the aspect of fixing Domain = Panjab University. Queries using six different keywords: person, faculty, expert, specialized, expertise and research were executed to find the expert(s) in the Field of “pattern recognition”, “phytoremediation” and “software engineering” in Panjab University.

Field “pattern recognition”

The queries formed using the Field “pattern recognition”, with fixing Domain = Panjab University retrieved more relevant results (RDJR) as compared to without fixing Domain.

- Precision (P) and Recall (R) for the keywords: person, expert and specialized remained zero both with and without fixing Domain.
- Keywords: person, expert and specialized could not match the Concept of the problem.
- On fixing Domain, Precision and Recall for the keyword: faculty improved significantly with Recall = 0.5 and Precision = 0.1667.
- Precision for the keywords: expertise and research also improved after Domain fixing.
- For the keywords: expertise and research, Recall = 1. These are the most suitable keywords to retrieve maximum number of relevant results.
- This concludes improved performance after fixing Domain.

Field “phytoremediation”

The queries formed using the Field “phytoremediation”, with fixing Domain = Panjab University retrieved more relevant results (RDJR) as compared to without fixing Domain.
- Precision and Recall for the keywords: person, expert and specialized remained zero both with and without fixing Domain.
- Keywords: person, expert and specialized could not match the Concept of the problem.
- On fixing Domain, Precision for the keywords: faculty, expertise and research improved significantly.
  - Precision (faculty) = 0.25
  - Precision (expertise) = 1.0
  - Precision (research) = 0.1667
- There was no change in the Recall for these keywords after fixing Domain.
- After fixing Domain, highest Precision was obtained for keyword: expertise.
- This concludes improved performance after fixing Domain.

Field “software engineering”

The queries formed using the Field “software engineering” with fixing Domain = Panjab University retrieved more relevant results (RDJR) as compared to without fixing Domain.
- Precision and Recall for the keywords: person, expert and specialized remained zero both with and without fixing Domain.
- Keywords: person, expert and specialized could not match the Concept of the problem.
- On fixing Domain, Precision for the keywords: faculty, expertise and research improved significantly.
On fixing Domain, Recall for the keywords: faculty and research also improved.

- There was no change in the Recall for the keyword: expertise. It remained at the highest value = 1 both without and with fixing Domain.
- After fixing Domain, highest Precision was obtained for keyword: expertise.
- This concludes improved performance after fixing Domain.

**Metrics used for the evaluation of Noise Removal Framework was Mean Retrieved Documents Judged Relevant (RDJR).** Descriptive Statistics: Mean and Inferential Statistics: Independent t-test are the statistical tools used for analysis.

**Hence Proved 1:** The result of t-test shows that there is significant difference between the Mean value of RDJR for the three fields “pattern recognition”, “phytoremediation” and “software engineering” without fixing Domain = Panjab University and with fixing Domain = Panjab University.

**Hence Proved 2:** Mean RDJR with fixing is more as compared to without fixing Domain. So, more number of retrieved documents were judged as relevant after fixing Domain.

This proves the efficiency of the suggested Noise Removal Framework from the aspect of fixing Domain = Panjab University especially in terms of Precision i.e. the ability to retrieve top-ranked results that are mostly relevant.

### 6.2 Noise Removal for Semantic information Processing (NRSIP) Framework (117)

The retrieval effectiveness of NRSIP framework is discussed in comparison to PSE (Yahoo!) for the following five parameters:

- **Query Retrieval Time (QRT):** Significant difference in the Mean QRT (PSE) as compared to Mean QRT (NRSIP) has been found. More time (QRT) has been taken by PSE to retrieve the query results as compared to NRSIP framework.
NRSIP framework substantially improves the QRT.

- **Precision (P):** Significant difference in the Precision (PSE) and Precision (NRSIP) has been found. Mean for Precision (NRSIP) is more as compared to Mean for Precision (PSE), so more number of Retrieved Documents were Judged as Relevant by NRSIP framework as compared to PSE (Table 5.14 and Figure 5.8). Hence, NRSIP framework is more efficient in terms of Precision. NRSIP framework substantially improves the Precision.

- **Proportion of Retrieved Documents Judged Irrelevant (RDJI):** Significant difference in the Proportion of RDJI (PSE) and Proportion of RDJI (NRSIP) has been found. Mean for Proportion of RDJI (PSE) is more as compared to Mean for Proportion of RDJI (NRSIP), so more number of Irrelevant Documents were Retrieved by PSE as compared to NRSIP framework (Table 5.15 and Figure 5.9). This proves the efficiency of NRSIP framework over PSE in terms of RDJI. NRSIP framework substantially improves the Proportion of RDJI.

- **Recall (R):** Significant difference in the Recall (PSE) and Recall (NRSIP) has been found. Mean for Recall (NRSIP) is more as compared to Mean for Recall (PSE), so more number of Relevant Documents were Retrieved by NRSIP framework as compared to PSE (Table 5.16 and Figure 5.10). Hence, NRSIP framework is more efficient in terms of Recall. NRSIP framework substantially improves the Recall.

- **Proportion of Relevant Documents Missed (RDM):** Significant difference in the Proportion of RDM (PSE) and Proportion of RDM (NRSIP) has been found. Mean for Proportion of RDM (PSE) is more as compared to Mean for Proportion of RDM (NRSIP), so more number of Relevant Documents were Missed by PSE as compared to NRSIP framework (Table 5.17, Figure 5.11). This depicts that NRSIP framework is more efficient over PSE in terms of RDM. NRSIP framework substantially improves the Proportion of RDM.

Overall, NRSIP framework outperforms PSE in all the above-mentioned retrieval effectiveness measures. This shows the success of the proposed and implemented NRSIP framework.
6.3 Noise Removal for Semantic Information Processing (NRSIP) Framework for Extended Dataset

The retrieval effectiveness of NRSIP framework for extended dataset is discussed in comparison to PSE (Yahoo!) for the following parameters:

- **Query Retrieval Time (QRT):** Significant difference in the Mean QRT (PSE) as compared to Mean QRT (NRSIP) has been found. More time (QRT) has been taken by PSE to retrieve the query results as compared to NRSIP framework for extended dataset (Table 5.49 and Figure 5.13). Hence, PSE is less efficient in terms of QRT. **NRSIP framework substantially improves the QRT for extended dataset.**

- **Precision (P):** Significant difference in the Precision (PSE) and Precision (NRSIP) has been found. Mean for Precision (NRSIP) is more as compared to Mean for Precision (PSE), so **more number of Retrieved Documents were Judged as Relevant by NRSIP framework for extended dataset as compared to PSE** (Table 5.51 and Figure 5.14). Hence, NRSIP framework is more efficient in terms of Precision for extended dataset. **NRSIP framework substantially improves the Precision for extended dataset.**

- **Proportion of Retrieved Documents Judged Irrelevant (RDJI):** Significant difference in the Proportion of RDJI (PSE) and Proportion of RDJI (NRSIP) has been found. Mean for Proportion of RDJI (PSE) is more as compared to Mean for Proportion of RDJI (NRSIP), so **more number of Irrelevant Documents were Retrieved by PSE as compared to NRSIP framework for extended dataset** (Table 5.52). Thus, NRSIP framework substantially improves the Proportion of RDJI for extended dataset.

- **Recall (R):** Significant difference in the Recall (PSE) and Recall (NRSIP) has been found. Mean for Recall (NRSIP) is more as compared to Mean for Recall (PSE), so **more number of Relevant Documents were retrieved by NRSIP framework as compared to PSE** (Table 5.53 and Figure 5.15). Hence, NRSIP framework substantially improves the Recall for extended dataset.

Overall, NRSIP framework outperforms PSE for extended dataset. This shows the success of NRSIP framework for extended dataset.
6.4 Generalized Noise Removal for Semantic Information Processing Framework (NRSIP-G)

The retrieval effectiveness of NRSIP-G framework is discussed in comparison to PSE (Yahoo!) for the following parameters:

- **Query Retrieval Time (QRT):** Significant difference in the Mean QRT (PSE) as compared to Mean QRT (NRSIP-G) has been found. More time (QRT) has been taken by PSE to retrieve the query results as compared to NRSIP-G framework (Table 5.85 and Figure 5.17). Hence, PSE is less efficient in terms of QRT. **NRSIP-G framework substantially improves the QRT.**

- **Precision (P):** Significant difference in the Precision (PSE) and Precision (NRSIP-G) has been found. Mean for Precision (NRSIP-G) is more as compared to Mean for Precision (PSE), so more number of Retrieved Documents were Judged as Relevant by NRSIP-G framework as compared to PSE (Table 5.87 and Figure 5.18). Hence, NRSIP-G framework is more efficient in terms of Precision. **NRSIP framework substantially improves the Precision.**

Overall, NRSIP-G framework outperforms PSE in all the above-mentioned retrieval effectiveness measures. This shows the success of the proposed and implemented NRSIP-G framework.

6.5 Secured Cognitive Agent based Multi-strategic Intelligent Search System (CAbMsISS) (119)

The retrieval effectiveness of Secured CAbMsISS is discussed in comparison to PSE (Google) for the following three parameters:

- **Query Retrieval Time (QRT):** Significant difference in the Mean (QRT) for PSE and CAbMsISS has been found. More time has been taken by PSE to retrieve the query results as compared to CAbMsISS (Table 5.89 and Figure 5.19). Hence, PSE is less efficient in terms of QRT. **Therefore, CAbMsISS substantially improves the QRT.**

- **Precision (P):** Significant difference in the Precision (PSE) and Precision (CAbMsISS) has been found. Mean for Precision (CAbMsISS) is more as
compared to Mean for Precision (PSE), so more number of Retrieved Documents were Judged as Relevant by CAbMsISS as compared to PSE (Table 5.91 and Figure 5.20). Hence, CAbMsISS is more efficient in terms of Precision. Therefore, CAbMsISS substantially improves the Precision.

➢ Proportion of Retrieved Documents Judged Irrelevant (RDJI): Significant difference in the Proportion of RDJI (PSE) and Proportion of RDJI (CAbMsISS) framework has been found. Mean for Proportion of RDJI (PSE) is more as compared to Mean for Proportion of RDJI (CAbMsISS) so, more number of Irrelevant Documents were Retrieved by PSE as compared to CAbMsISS (Table 5.92 and Figure 5.21). Thus, proving the efficiency of CAbMsISS over PSE. Therefore, CAbMsISS substantially improves the Proportion of RDJI.

As CAbMsISS outperforms PSE in all the above-mentioned retrieval effectiveness measures. Hence, this illustrates the success of the proposed and implemented framework for CAbMsISS.

All the implemented frameworks including Noise Removal Framework (116), NRSIP framework (117), NRSIP-G framework and Secured CAbMsISS (119) have been found to be more efficient as compared to PSE. Therefore, the ultimate aim of improving the user satisfaction for the retrieval results by mapping the Concept and Context of the user has been achieved by all the frameworks.