



UNIVERSITI  
TEKNOLOGI  
PETRONAS

# LITERATURE REVIEW CHAPTER

## BEST PRACTICES IN CRAFTING LITERATURE REVIEW CHAPTER

Dr Noureen Talpur  
Lecturer  
Computer & Information Sciences Department



# DR NOUREEN TALPUR

Lecturer at Universiti Teknologi PETRONAS (UTP)



## EDUCATION



PhD IT, Universiti Teknologi PETRONAS, Malaysia



MSc. IT, Universiti Tun Hussein Onn Malaysia, Malaysia



B.CS, Muhammad Ali Jinnah University, Pakistan

## EXPERIENCE

- Graduate Assistant - 5 years
- Teaching Experience - 4 years



## RESEARCH / EXPERTISE

- Artificial Intelligence and Machine Learning for Smart Communities
- Data Science (Big Data Analytics, Data quality assessment, data mining)
- Optimization, prediction, classification, and feature selection

## CONSULTANCY AND TRAINING

- Master the Basics of Big Data, AI, and Machine Learning for Tomorrow's Workforce (a fully funded program by KPT)
- Data Understanding & Data Quality Assessment (PETRONAS GTS - OPTIMUS)

## PROFESSIONAL BODY / MEMBERSHIP

Graduate Technologist, MBOT



## TEACHING EXPERIENCE

- Artificial Intelligence (using MATLAB)
- Structured Algorithm & Programming (using Python)
- Object Oriented Programming (using Python)
- System Architecture
- Big Data Analytics (using Hadoop & Map-reduce framework)
- Discrete Mathematics
- Statistics and Empirical Methods



## ACHIEVEMENTS



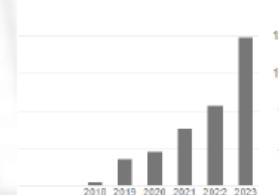
Cited by

All Since 2018

Citations 425 423

h-index 8 8

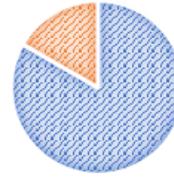
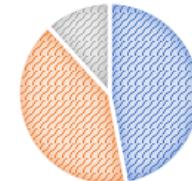
h10-index 7 7



## PUBLICATIONS (17)

- Journal Publications
- Conference Proceedings
- Book Chapters

Q1 Q2



Research Grants Utilized = 5



# OUTLINE

- 1. Purpose of a Literature Review**
- 2. Steps to search studies for Literature Review Chapter**
  - i. Search relevant literature aligned with your topic
  - ii. Analyze & evaluate the importance of paper
  - iii. Create an excel and list the important studies
- 3. Writing Literature Review**
  - i. Make an outline of the Chapter 2
  - ii. Ensure appropriate citation of literature
  - iii. Summarize literature studies in the form of table
  - iv. Addressing research gaps identified from existing literature
  - v. Conclusion remarks





## 1. Purpose of a Literature Review

- ✓ Literature review section of your thesis is where **you investigate past work relevant to your research topic.**
- ✓ Define **key concepts and ideas**
- ✓ Following Chapter 1, your literature review embarks on **answering your primary research question: Who** has explored **what** aspects, **why**, and **what** discoveries have they made?
- ✓ It **enables you to grasp the existing work done of your topic**, validate your **hypotheses**, **fine-tune** your research questions, and **pinpoint any existing gaps**.
- ✓ It also **serves as an evidence of your familiarity** with relevant research **within your field**.





## 2. Steps to search studies for Literature Review Chapter



- i      Search relevant literature aligned with your topic
- ii     Analyze & evaluate the importance of paper
- iii    Create an excel and list the important studies



UNIVERSITI  
TEKNOLOGI  
PETRONAS

## 2. Steps to search studies for Literature Review Chapter

### i - Search relevant literature aligned with your topic



Google Scholar

<https://scholar.google.com/>

Find most relevant studies  
from past 5 years

The screenshot shows a Google Scholar search results page. The search query 'feature selection for high dimensional' is entered in the search bar. The results are filtered to show 'About 17,900 results (0.09 sec)' from 'Articles'. The search interface includes filters for 'Any time', date range from '2019 — 2023', and sorting options like 'Sort by relevance' and 'Sort by date'. The results list three academic papers:

- [HTML] Benchmark for filter methods for feature selection in high-dimensional classification data** [HTML] sciencedirect.com  
A Bommert, X Sun, B Bischl, J Rahnenführer... - ... Statistics & Data Analysis, 2020 - Elsevier  
... feature selection techniques on high-dimensional data sets. We compare 22 filter methods from different toolboxes on 16 high-dimensional ... which methods select the features of a data ...  
☆ Save 99 Cite Cited by 562 Related articles All 11 versions Web of Science: 294
- Variable-size cooperative coevolutionary particle swarm optimization for feature selection on high-dimensional data** [PDF] ieee.org  
XF Song, Y Zhang, YN Guo, XY Sun... - IEEE Transactions on ..., 2020 - ieeexplore.ieee.org  
... PROPOSED FEATURE SELECTION ALGORITHM Focusing on the challenge of the curse of dimensionality caused by high-dimensional features, we present a new FS method based ...  
☆ Save 99 Cite Cited by 264 Related articles All 2 versions Web of Science: 210
- A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction** [PDF] jastt.org  
R Zebari, A Abdulazeez, D Zeebaree, D Zebari... - Journal of Applied ..., 2020 - jastt.org  
... This work reviewed more than forty articles of feature selection and feature extraction that ...  
Various issues can benefit from the feature selection techniques application. High dimension, ...  
☆ Save 99 Cite Cited by 622 Related articles All 8 versions



## **ii - Analyze & evaluate the importance of paper**



# Benchmark for filter methods for feature selection in high-dimensional classification data

**1 - Read Abstract and Conclusion of the study to evaluate if the study closely match with your area of the study**

Under a Creative Commons license ↗

## 4. Summary and conclusions

Feature selection is a key part of [data analysis](#), machine learning and [data mining](#). There are many methods for feature selection, but it is unclear which of these methods perform best. In this analysis we focused on the comparison of filter methods for feature selection. We analyzed 22 filter methods based on 16 high-dimensional classification data sets from various domains.

To find out which of the filter methods are similar, we first sorted them by rank. Then we rank the features, we computed [rank correlation](#) between the feature rankings of all pairs of methods. For each method, we computed the average rank correlation with all other methods. For each data set, especially the data sets containing gene expression profiles, we found that there were three groups of similar methods. The methods in one group were not similar to any other method. The methods in another group were similar to each other and came from the same toolboxes. For the other methods, the similarity was not very clear. Also, we investigated the scaling behavior of the methods. We found that there were two groups of filters that behave similarly with respect to scaling.

## 2 - Down

- Universiti Teknologi PETRONAS > Desktop > Literature Studies

Name
2019 - Feature selection based on artificial bee colony and gradient boosting decision tree
2020 - Mayfly_in_Harmony_A_New_Hybrid_Meta-Heuristic_Feature_Selection_Algorithm
2021 - Hybrid_Binary_Grey_Wolf_With_Harris_Hawks_Optimizer_for_Feature_Selection
2022 - Efficient high-dimension feature selection based on enhanced
2023 - A Novel Wrapper-Based Optimization Algorithm for the Feature Selection an.

**2 - Download the paper if its relevant to read details later**

**3 - Create a folder and save all those studies  
organized by year of study**



# 1. Steps to search studies for Literature Review Chapter

UNIVERSITI  
TEKNOLOGI  
PETRONAS

## iii - Create an excel and list of the important studies



IEEE  
**Xplore**<sup>®</sup>  
Digital Library

<https://ieeexplore.ieee.org/Xplore/home.jsp>

The screenshot shows the IEEE Xplore Digital Library interface. A search has been conducted for "feature selection for high dimensional". The results page displays 1-25 of 4,318 results. The "Export" button is highlighted with a yellow box and an orange arrow. A modal window titled "Download Results" shows that 25 results have been selected for download to a CSV file. The "Download" button in this modal is also highlighted with a yellow box and an orange arrow. In the bottom right corner, a Microsoft Excel spreadsheet titled "export2024.06.05-03.51.25" is shown, containing a table of search results with columns for Document Authors, Author Affiliation, Publication Date, Add Date, Publication Start Page, End Page, Abstract, and ISSN.

Document Authors	Author Affiliation	Publication Date	Add Date	Publication Start Page	End Page	Abstract	ISSN
A Modified M. Xuan; L. College of	2021 IEEE	14-Apr-22	2021	109	113	Feature selection (F)	
A Hybrid S. U. Ali; K. S. College of	2019 Interi	21-Jun-19	2019	71	75	Now a days, in educ	
Feature Se G. Manikai Anna Univ	2017 Seco	5-Oct-17	2017	320	325	In recent years, feat	
Flexible-Fi G. Manikai Departme	2018 Tenti	#####	2018	237	243	Due to advancement	
An effectiv S. L. Happy Departme	2017 25th	#####	2017	1574	1578	Feature se 2076-14	
Extremely S. Li; D. Wi College of	IEEE Trans	#####	2014	737	747	To select ii 2168-22	
A hybrid fe A. Rouhi; F. Departme	2017 2nd C	8-Jun-17	2017	16	20	Nowadays, with the	
Parallel Lo Z. Manbari Departme	2021 11th	#####	2021	500	507	Recent tec 2643-27	
Preordona H. Chamla Departme	2021 Fifth	1-Dec-21	2021	1	5	Feature selection is	
Research (M. Zhu; B. Zhejiang U	2017 Interi	#####	2017	273	277	The purpose of this	

1. **Search by topic**
2. **Choose by Conference, journal, book, magazine ...**
3. **Select the studies matching to your title/research area**
4. **Export the list in CSV**



## 2. Steps to search studies for Literature Review Chapter

UNIVERSITI  
TEKNOLOGI  
PETRONAS

### iii - Create an excel and list of the important studies

*Create your own Excel file of the studies that are most relevant, and you want to add into your thesis*

A	B	C	D	E	F	G
Author and Year	Techniques	Application	Contribution	Results	Comparison with	Outcome
Zheng, et al., 2016	deep Boltzmann machine (DBM) with Pythagorean fuzzy (PFDBM)	Airline Passenger Profiling	i) Pythagorean-type fuzzy sets to improve the representation ability and robustness. ii) hybrid-training algorithm is using the greedy gradient-based method for efficiently locating an optimum in a small region of the solution space. iii) BBO to explore a wider area to avoid being trapped in local optima. iv) The hybrid algorithm applies to the PFDBM layer by layer.	DBM, FDBM, PFDBM-G, PFDBM-BBO		The final results concluded that the models PFDBM-G with 57.14%, and PFDBM-BBO with 61.90% received better precision than DBM and FDBM with 30.95% and 35.71% of the precision. While the proposed model of PFDBM-EBO outperformed all four models with 64.29% of precision.
Zheng, et al., 2017	Pythagorean-type fuzzy with deep denoising autoencoder (PFDDAE)	Industrial Accident Early Warning	i) equips the classical deep denoising autoencoder (DDAE) model with Pythagorean-type fuzzy parameters in order to enhance the model's representation ability and robustness. ii) To efficiently train the fuzzy deep model, combines biogeography-based optimization (BBO) for facilitating exploration iii) Hessian-free (HF) optimization algorithm for enhancing exploitation in the solution space.	DDAE, FDDAE, PFDDAE-HF, PFDDAE-GA, PFDDAE-BBO		DDAE 20.63%, FDDAE 17.71%, PFDDAE-HF 16.54%, PFDDAE-GA 15.58%, PFDDAE-BBO 15.13%
Chimatapu, et al.,	Type-2 Fuzzy Logic	Workforce	i) A novel Type-2 Fuzzy Logic System (FLS) is	Standard SAE with 2 hidden		The Shallow Neural Network (MLP) with one hidden layer



### 3. Writing Literature Review

#### i - Make an outline of the Chapter 2

##### Introduction

- Overview **Purpose and Importance** of the Literature Review - Explanation of the **Structure and Organization** of the Chapter

##### Theoretical Framework

- **Introduction** to **Theoretical Frameworks** relevant to the research topic
- Explanation of **key concepts** and constructs

##### Related Works

- **Exploration of historical background** of Topic - Review of **recent studies** and developments in the field - **Comparison of different approaches and methodologies**

##### Research Gap

- Identify **research gaps/unanswered questions** in literature - Discuss **challenges/limitations** in past studies – **Justify why addressing those gaps are important**

##### Chapter Summary

- **Recap of key findings/insights** from the literature review - Summary of **significance of the literature review** to the overall study



### 3. Writing Literature Review

#### ii - Ensure appropriate citation of literature

- ✓ **Always cite works of other authors** whose ideas, research findings, theories, or arguments or **even Figures** you have included or discussed in your literature review. **Citing your resources** is required for all kinds of writing projects.
- ✓ It is **important to cite, not only direct quotes**, but **paraphrased sentences** as well. A paraphrased sentence is when you take a passage from a source and rewrite it in your own words.

Example of quote: ***Evolution is a gradual process that “can act only by very short and slow steps” (Darwin, 1859).***

- ✓ **Maintain only one citation style** (APA, MLA, Chicago, IEEE) throughout your literature review.

<https://www.sjsu.edu/writingcenter/>

Popular Citation Styles	
And Who Uses Them!	
	<b>MLA</b> English Literature Foreign Language Communications Religious Studies
	<b>APA</b> Psychology Education Business/Economics Nursing Linguistic Studies
	<b>Chicago</b> History Art History Fine Arts Anthropology Philosophy
	<b>IEEE</b> Engineering Computer Science Information Science
	<b>CSE</b> Biology Physics Chemistry Geology

Always check with your instructor about what style to use!





### 3. Writing Literature Review

#### ii - Ensure appropriate citation of literature

#### Helpful Phrases for Writing Literature Review

##### 1. Previous literature

The literature review shows that \_\_

Previous research showed \_\_

Seminal contributions have been made by \_\_

A series of recent studies has indicated that \_\_

Several theories have been proposed to \_\_, some focusing on \_\_, others on \_\_

There has been numerous studies to investigate \_\_

This has been used in several studies to assess \_\_

Previous studies have shown \_\_

Several studies suggest that \_\_

This has also been explored in prior studies by \_\_

Prior research suggests that \_\_

Previous studies have emphasized \_\_

The majority of prior research has applied \_\_

Most early studies as well as current work focus on \_\_

For instance, the following studies were conducted on \_\_

Studies of \_\_are well documented, it is also well acknowledged that \_\_

A number of authors have recognized \_\_

##### 2. Limitations of previous research

A number of questions regarding \_\_ remain to be addressed.

A closer look to the literature on \_\_, however, reveals a number of gaps and shortcomings.

This question has previously never been addressed because \_\_

Most studies have relied on \_\_

Previous studies by \_\_ cannot be considered as conclusive because \_\_

Previous studies have almost exclusively focused on \_\_

This has been previously assessed only to a very limited extent because \_\_

In the present studies \_\_ were constrained to \_\_

In previous studies were limited to \_\_

Although results appear consistent with prior research, they appear inconsistent with \_\_

These are previously unstudied because \_\_

As far as we know, no previous research has investigated \_\_

Moreover, although research has illuminated \_\_ no study to date has examined \_\_

Despite decades of research, this continues to be debated among \_\_

This section points out some of the problems encountered in the extant research.

Although there are many studies, the research in \_\_ remains limited.



### 3. Writing Literature Review

#### ii - Ensure appropriate citation of literature

#### Helpful Phrases for Writing Literature Review

##### 3. Research questions

More specific research questions will be introduced and investigated in \_\_

A further question is whether \_\_

Finally, another promising line of research would be \_\_

The study addresses several further questions on \_\_

Some of the interesting questions in this context are \_\_

In order to address the questions outlined above, we report here \_\_

These questions are of central interest as much recent research in \_\_

Furthermore, \_\_ is arguably an important question to be addressed.

The question now is how \_\_ can be used to explain \_\_

Study addresses the research question \_\_

In order to properly address this question, we \_\_

An important question associated with \_\_ is \_\_

A critical open question is whether \_\_

A still unsolved question is whether \_\_

This remains an open question as \_\_

This question has previously never been addressed because \_\_

This study offers a test of \_\_ research question

Study addresses the research question \_\_

##### 4. Research to be explored

A more systematic and theoretical analysis is required for \_\_

As the authors note earlier, more work is necessary to \_\_

Additional studies to understand more completely the key tenets of \_\_ are required.

The unexpected findings signal the need for additional studies to understand more about \_\_

This paper addresses \_\_, so far lacking in the scientific literature.

A new approach is therefore needed for \_\_

One of the tough challenges for all researchers in this domain is \_\_



# 1. Writing Literature Review

## iii - Summarize literature studies in the form of table

**Table 2:** Summary of the recent literature for the task of FS in medical sciences

Study, year	Algorithm	Application	Major contributions	Outcomes
[61], 2019	Crow-Search Algorithm (CSO)	Lung cancer detection	<ul style="list-style-type: none"><li>• Lungs modality is determined using computed tomography (CT) imaging, and classification is done using Probabilistic Neural Networks (PNN).</li><li>• Feature extraction is done using the gray-level co-occurrence matrix.</li><li>• CSO is used for the task of FS</li></ul>	The findings show that the CSO-based FS technique successfully delivered 90% accuracy.
[62], 2020	Two-stage Genetic Algorithm-feature Transformation (2-Tra-GA)	Hepatitis prediction	<ul style="list-style-type: none"><li>• A 2-Tra-GA-based feature transformation is proposed with two phases of FS and augmentation</li><li>• The lagrangian SVM (LSVM) approach is used to predict the hepatitis</li></ul>	The novel 2-Tra-GA-LSVM technique showed faster convergence with 90.32% accuracy in distinguishing between survived and deceased patients with hepatitis.
[63], 2021	Krill Herd Algorithm (KHA)	Brain epilepsy seizure detection	<ul style="list-style-type: none"><li>• The EEG data is used to identify epileptic seizures using a deep learning approach.</li><li>• KHA identifies the most optimal features from the brain signal data.</li><li>• General Adversarial Networks classify the selected features.</li></ul>	The KNN classifier detects the seizure with 99.25% when compared to the baseline method owing to the efficient FS mechanisms of the KHA algorithm.
[64], 2021	Grasshopper Optimization Algorithm (GOA)	Diabetes disease diagnosis	<ul style="list-style-type: none"><li>• GOA is used to select the key features from diabetes type II data.</li><li>• Three classifiers including SVM, NB, and Tree are used to classify the features of the diabetes data.</li></ul>	Compared to the NB and Tree classifier, SVM-GOA showed the best results having the least mean squared error of 3.0351.



### 3. Writing Literature Review

#### iii - Addressing research gaps identified from existing literature



- ✓ Recognizing areas of research having *insufficient/limited knowledge, understanding, or investigation*.
- ✓ These gaps can be identified by *unanswered questions* or *areas where further exploration is needed* such as Future works.
- ✓ Addressing these gaps involves *proposing potential research questions, methodologies, or approaches* that can contribute to filling these voids in knowledge.



### 3. Writing Literature Review

#### iii - Conclusion remarks



- ✓ In the conclusion of Chapter 2, the researcher should **recap the significant findings and insights** collected from the literature review.
  
- ✓ Highlight the importance of the literature review to the overall study, explaining its role in ***shaping the research framework***, ***guiding methodology***, and ***addressing gaps*** in the existing body of knowledge.



UNIVERSITI  
TEKNOLOGI  
PETRONAS

# Thank You!

