# Systematic Literature Review (SLR)

Romi Satria Wahono
*romi@romisatriawahono.net*
*http://romisatriawahono.net*
*08118228331*

# Romi Satria Wahono

- SMA Taruna Nusantara Magelang (1993)
- B.Eng, M.Eng and Ph.D in Software Engineering
  Saitama University Japan (1994-2004)
  Universiti Teknikal Malaysia Melaka (2014)
- Core Competency in Enterprise Architecture, Software Engineering and Machine Learning
- LIPI Researcher (2004-2007)
- Founder and CEO:
  - PT Brainmatics Cipta Informatika (2005)
  - PT IlmuKomputerCom Braindevs Sistema (2014)
- Professional Member of IEEE, ACM and PMI
- IT and Research Award Winners from WSIS (United Nations), Kemdikbud, Ristekdikti, LIPI, etc
- SCOPUS/ISI Indexed Journal Reviewer: Information and Software Technology, Journal of Systems and Software, Software: Practice and Experience, etc
- Industrial IT Certifications: TOGAF, ITIL, CCAI, CCNA, etc
- Enterprise Architecture Consultant: KPK, RistekDikti, INSW, BPPT, Kemsos Kemenkeu (Itjend, DJBC, DJPK), Telkom, FIF, PLN, PJB, Pertamina EP, etc

RomiSatriaWahono.Net

*Lecture Notes in Software Engineering, Computing Research and Technopreneurship*

HOME ABOUT ME IN T

BAGAIMANA MELAKU

Pada artikel ini, saya men
penelitian yang ditulis ole
to do Good Research" ini,
artikel yang saya tulis di b

TAHAPAN MEMULAI PENELITIAN UNTUK MAHASISWA GALAU

Jujur, secara umum saya
kecewa dengan pertanyaa
mahasiswa tingkat akhir y
masuk lewat email, inbox

RomiSatriaWahono.Net

*Lecture Notes in Software Engineering, Computing Research and Technopreneurship*

HOME ABOUT ME

Home » Lectures

SHARE THIS


Romi Satria Wahono

4.52K subscribers

SUBSCRIBED

HOME VIDEOS PLAYLISTS COMMUNITY CHANNELS ABOUT

Uploads PLAY ALL

SORT BY


5 Karakter Para Inovator: Pola Karakter yang Bisa...

586 views • 5 days ago


Knowledge Management (Pengelolaan Pengetahuan)...

934 views • 2 weeks ago


Systematic Literature Review (SLR): Jadi Nggak Galau La...

1.7K views • 3 weeks ago


Analisis dan Desain Software berbasis Unified Modeling...

916 views • 3 weeks ago


Software Engineering Research Trends: Analisis...

712 views • 4 weeks ago


Research Methodology (Metodologi Penelitian):...

1.2K views • 1 month ago


10 Mitos Kesalahan Penelitian Computing yang...

1.2K views • 1 month ago


10 Mitos Software Engineering yang Membuat...

1.1K views • 1 month ago


Praktek Cepat dan Mudah Data Mining dengan...

1.3K views • 1 month ago


Konsep, Proses dan Contoh Nyata Data Mining dan Big...

2.8K views • 1 month ago

# Course Outline

1. Pengantar Penelitian
   - 1.1 Konsep Penelitian
   - 1.2 Klasifikasi Penelitian
2. Literature Review
   - 2.1 Tahapan Penelitian
   - 2.2 Literature Review
3. Systematic Literature Review (SLR)
   - 3.1 Tahapan SLR
   - 3.2 Contoh SLR

# 1. Pengantar Penelitian

1.1 Konsep Penelitian

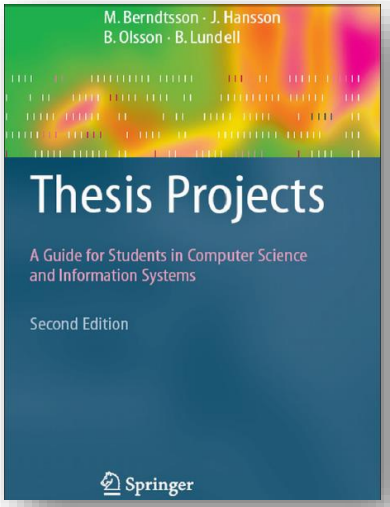
1.2 Klasifikasi Penelitian

# 1.1 Konsep Penelitian

# Mengapa Melakukan Penelitian?

- Berangkat dari adanya masalah penelitian
  - yang mungkin sudah diketahui metode pemecahannya
  - tapi belum diketahui metode pemecahan yang lebih baik

- Research (Inggris) dan recherche (Prancis)
  - re (kembali)
  - to search (mencari)



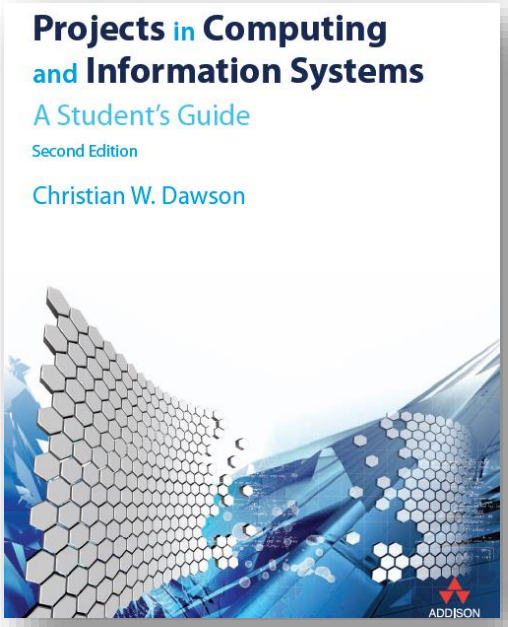

- The process of exploring the unknown, studying and learning new things, building new knowledge about things that no one has understood before
  *(Berndtsson et al., 2008)*

# Apa Yang Dikejar di Penelitian?

Research is a considered activity, which aims to make an original contribution to knowledge

*(contribution to the body of knowledge, in the research field of interest)*

*(Dawson, 2009)*

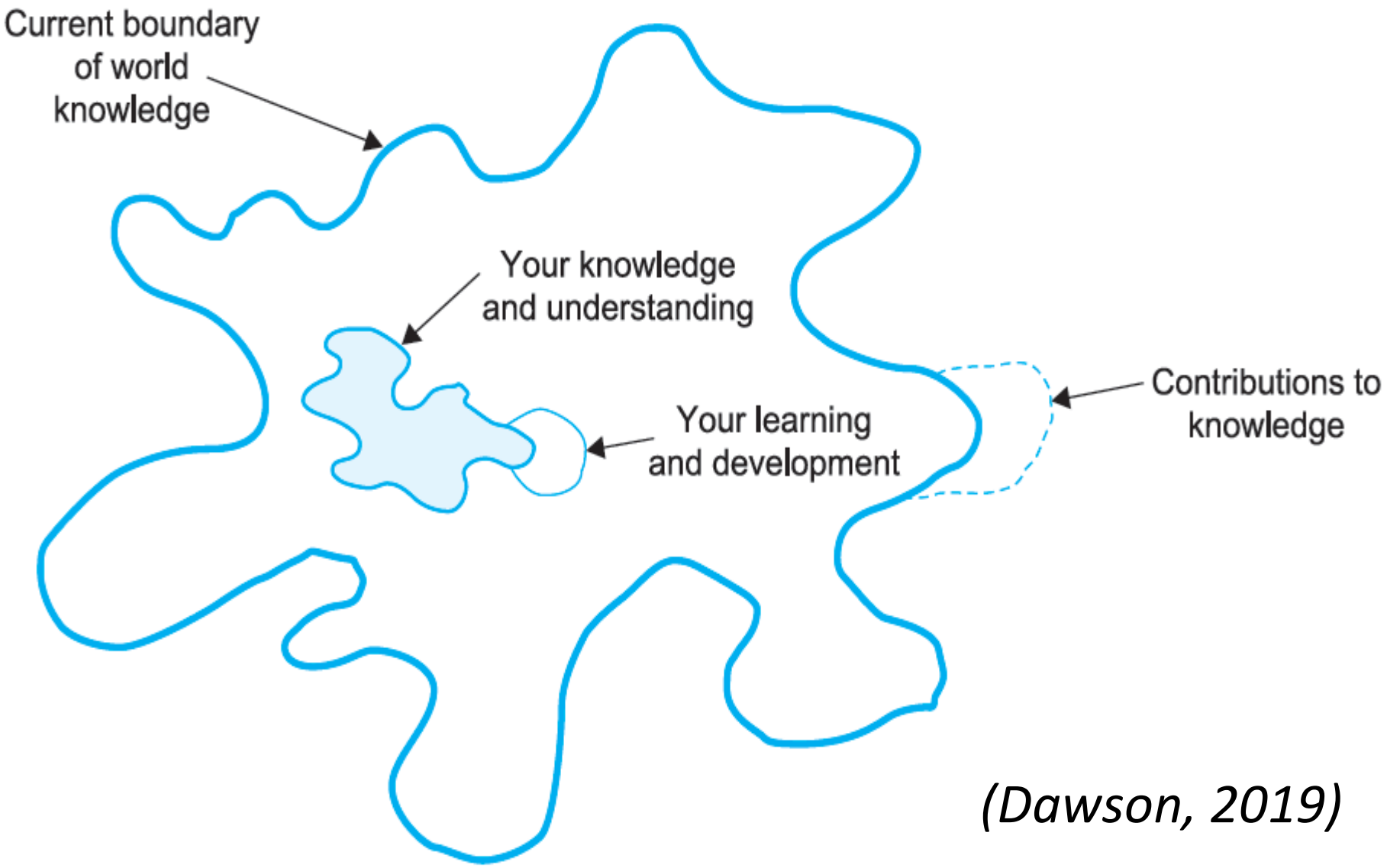



# Bentuk Kontribusi ke Pengetahuan

Kegiatan penyelidikan dan investigasi terhadap suatu masalah yang dilakukan secara berulang-ulang dan sistematis, dengan tujuan untuk menemukan atau merevisi teori, metode, fakta, dan aplikasi

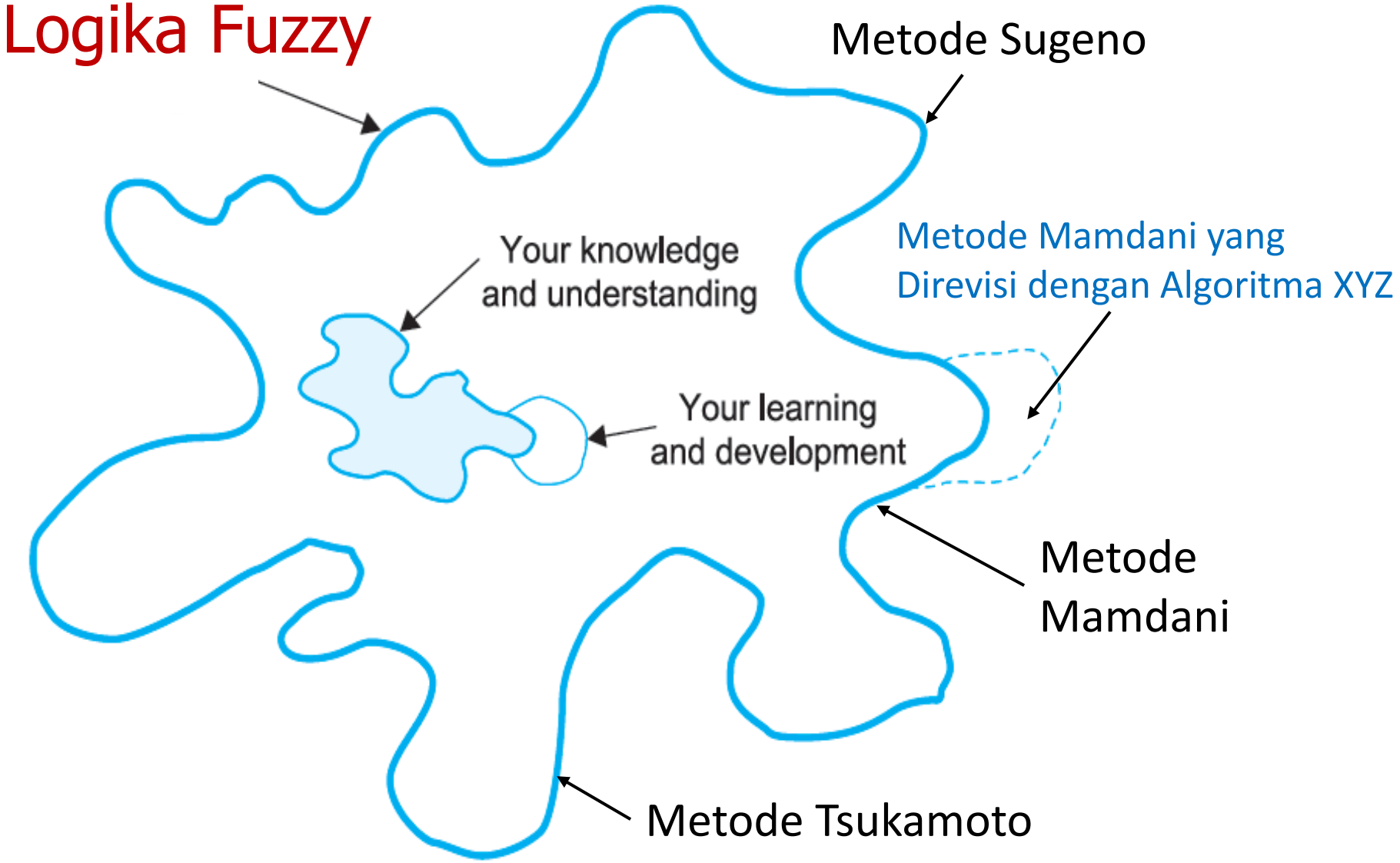
*(Berndtsson et al., 2008)*

# Bentuk Kontribusi ke Pengetahuan

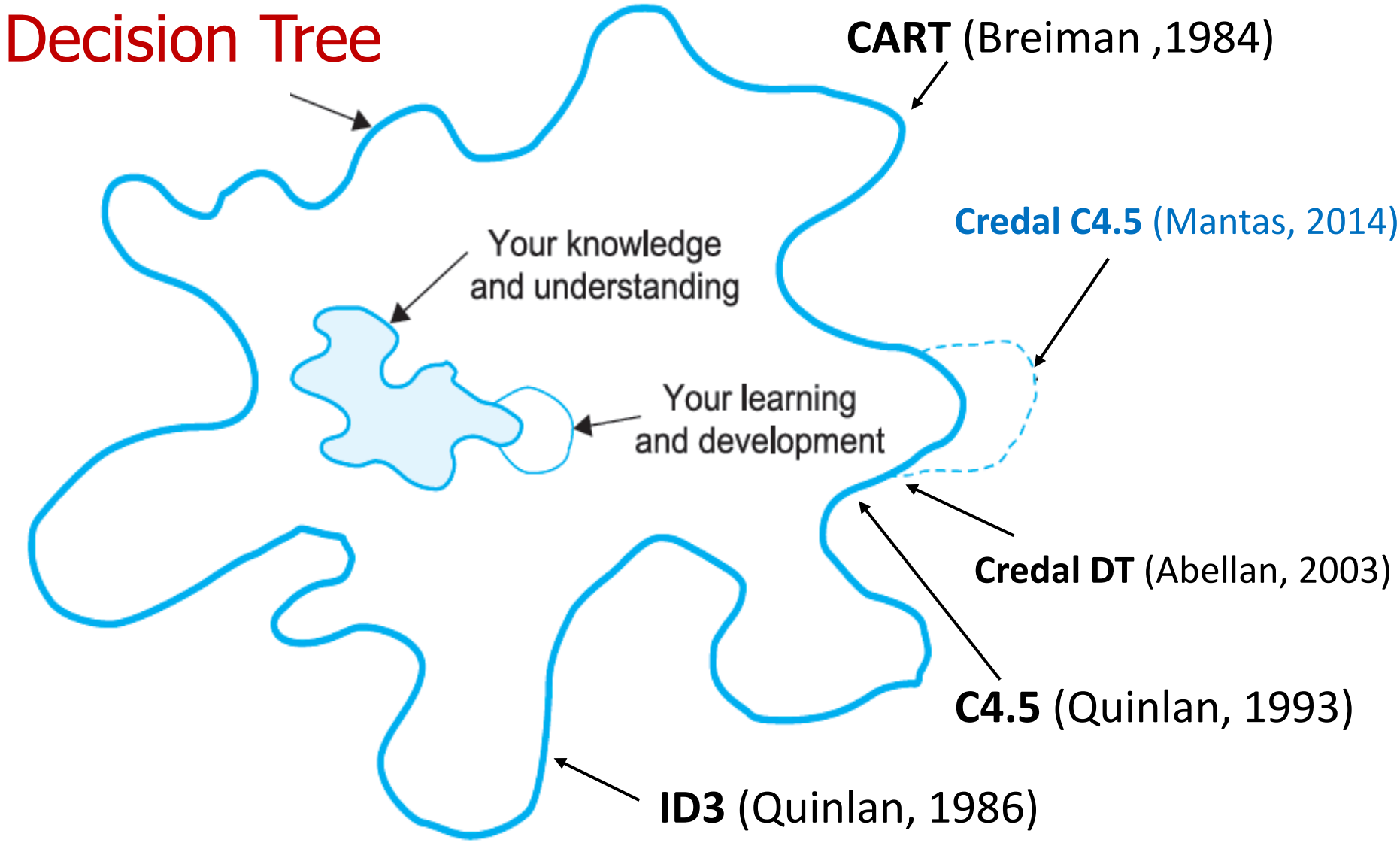



*(Dawson, 2019)*

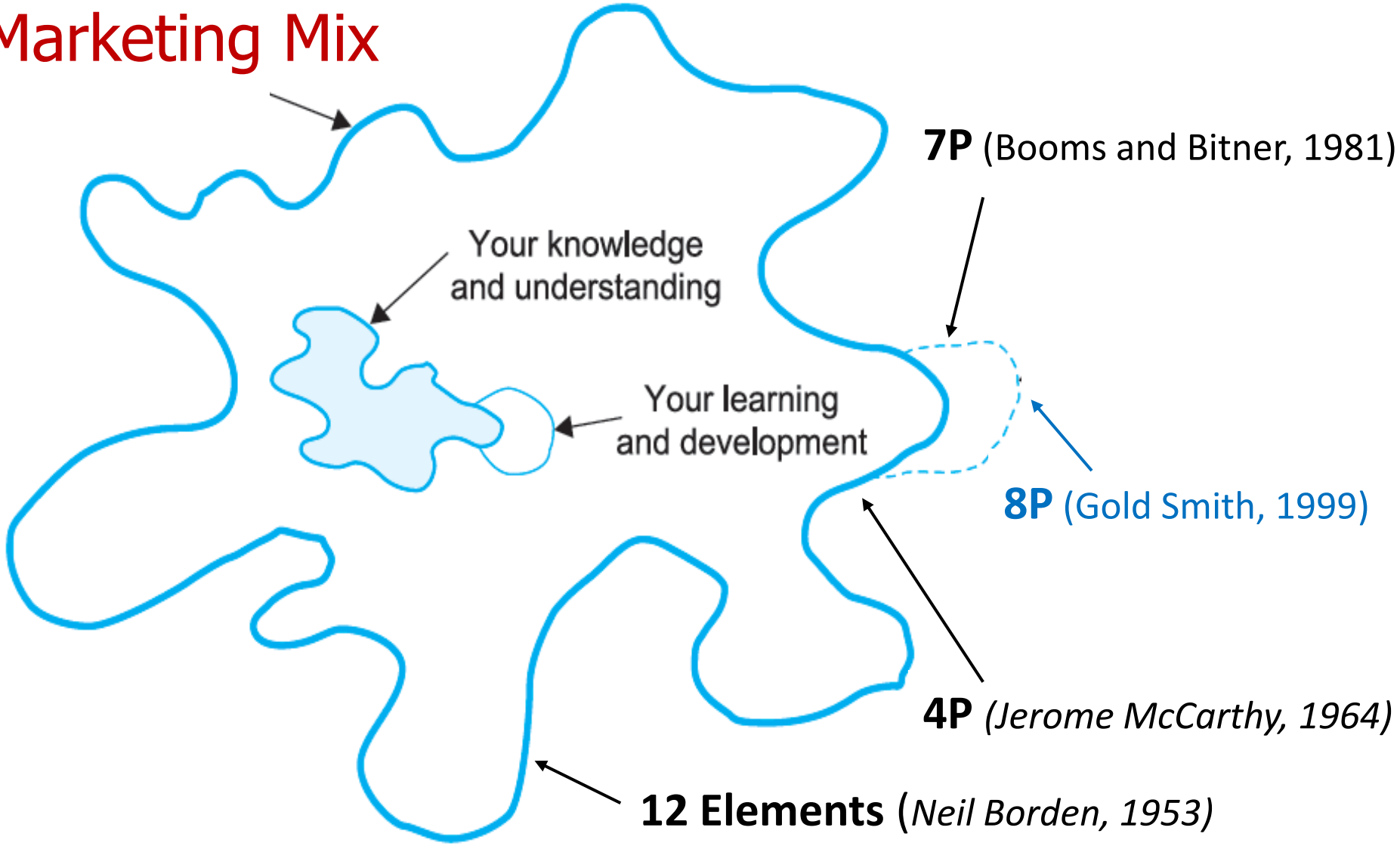
# Bentuk Kontribusi ke Pengetahuan



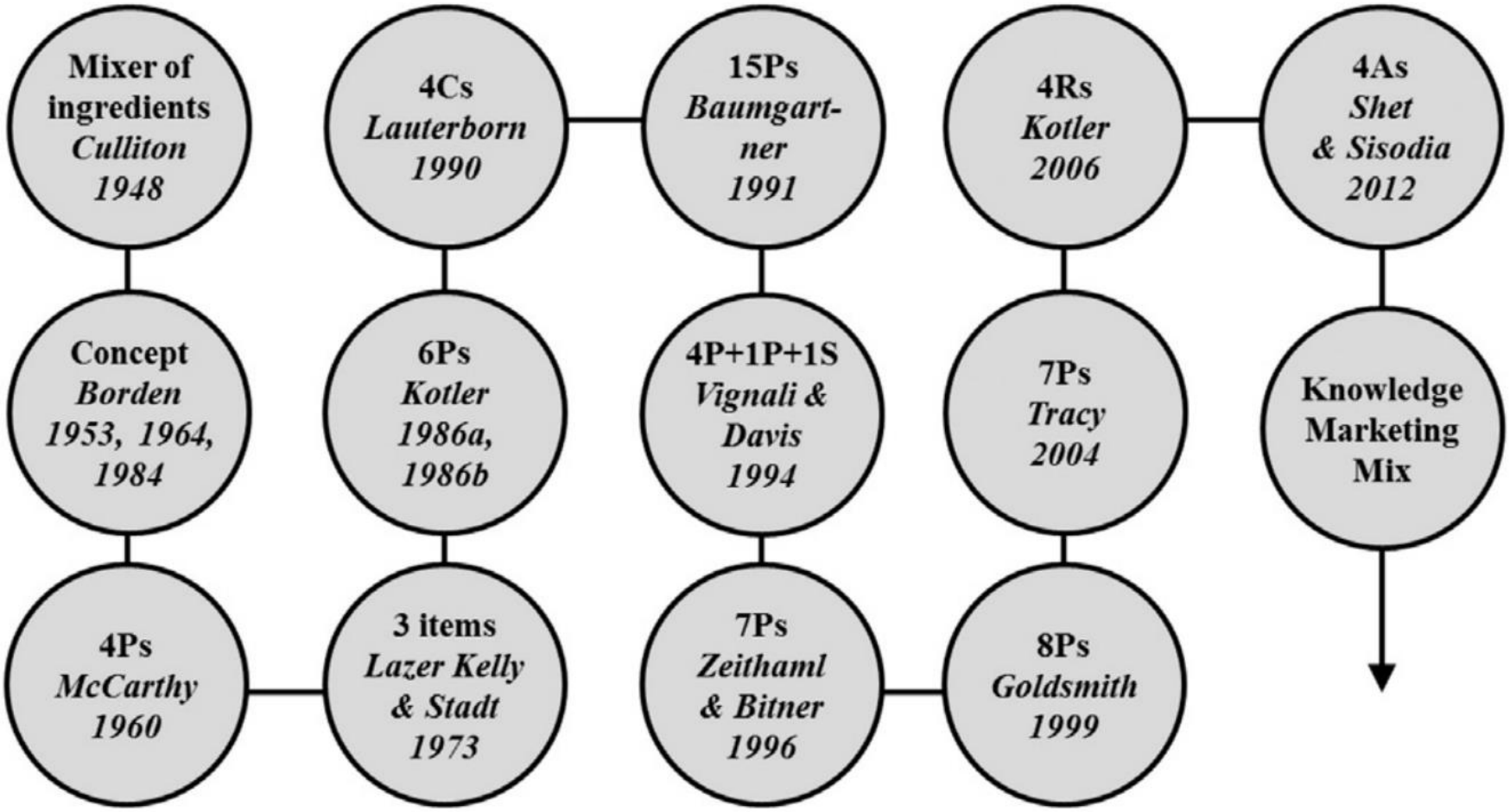
# Bentuk Kontribusi ke Pengetahuan

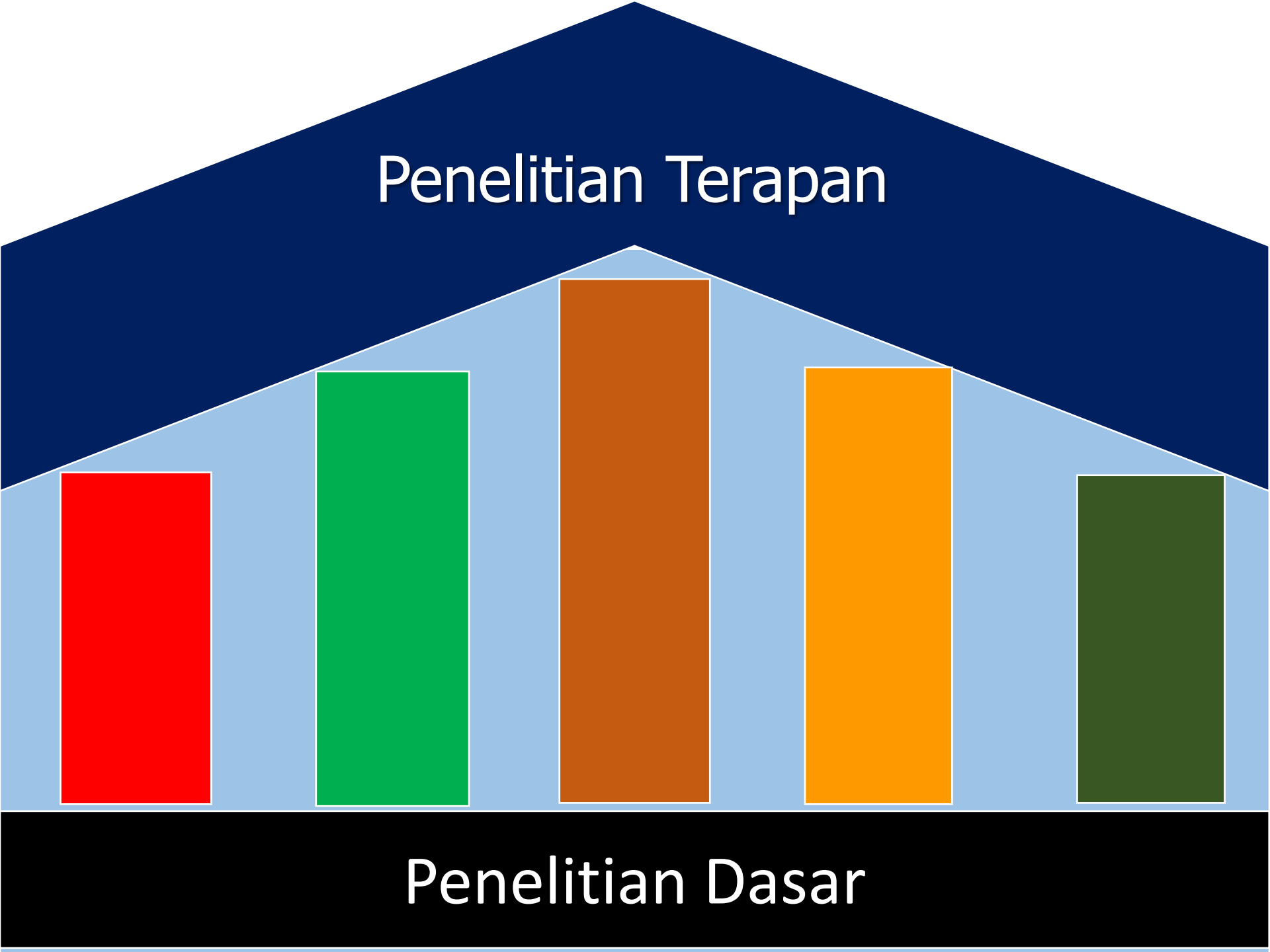


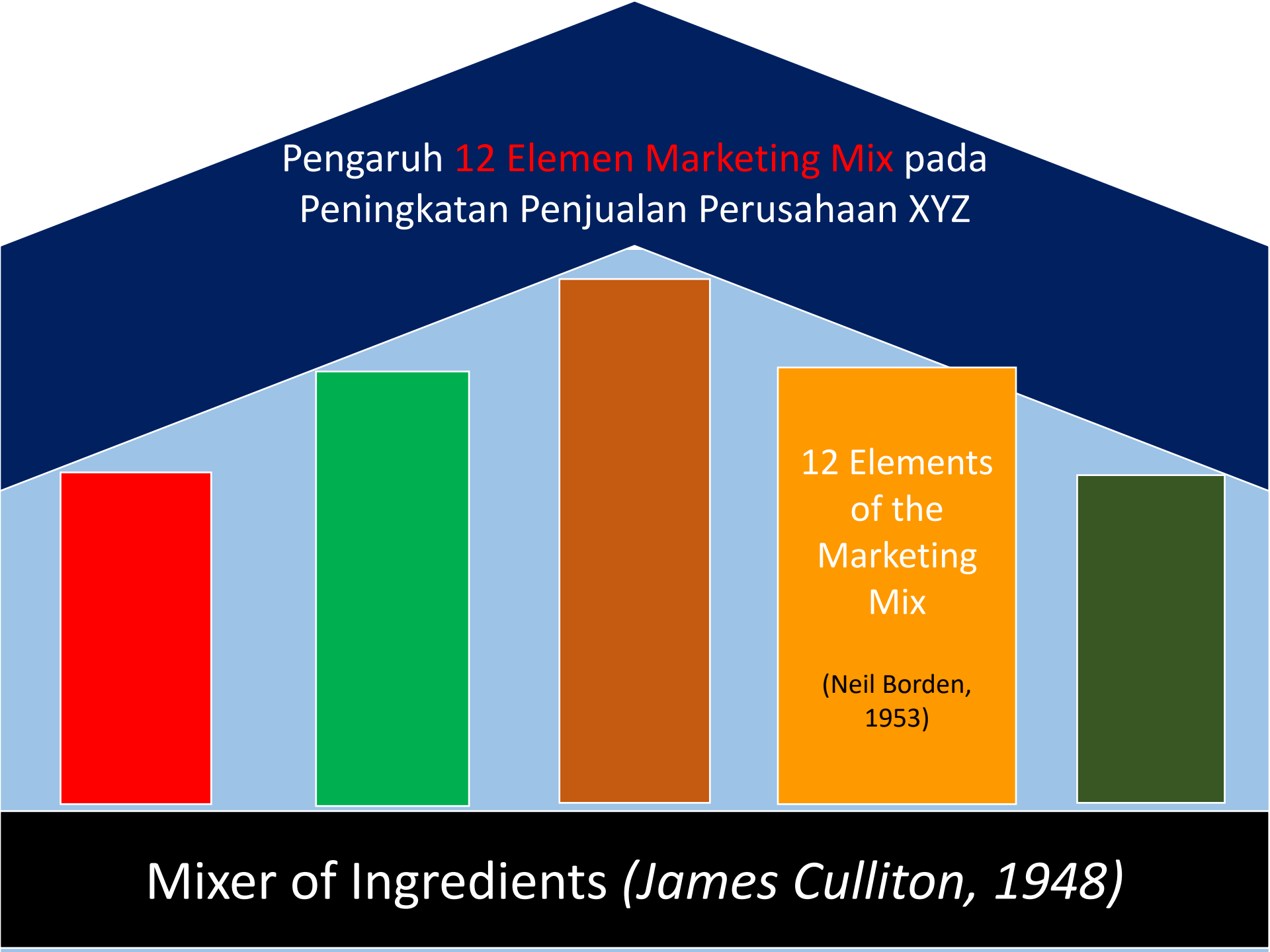
# Bentuk Kontribusi ke Pengetahuan

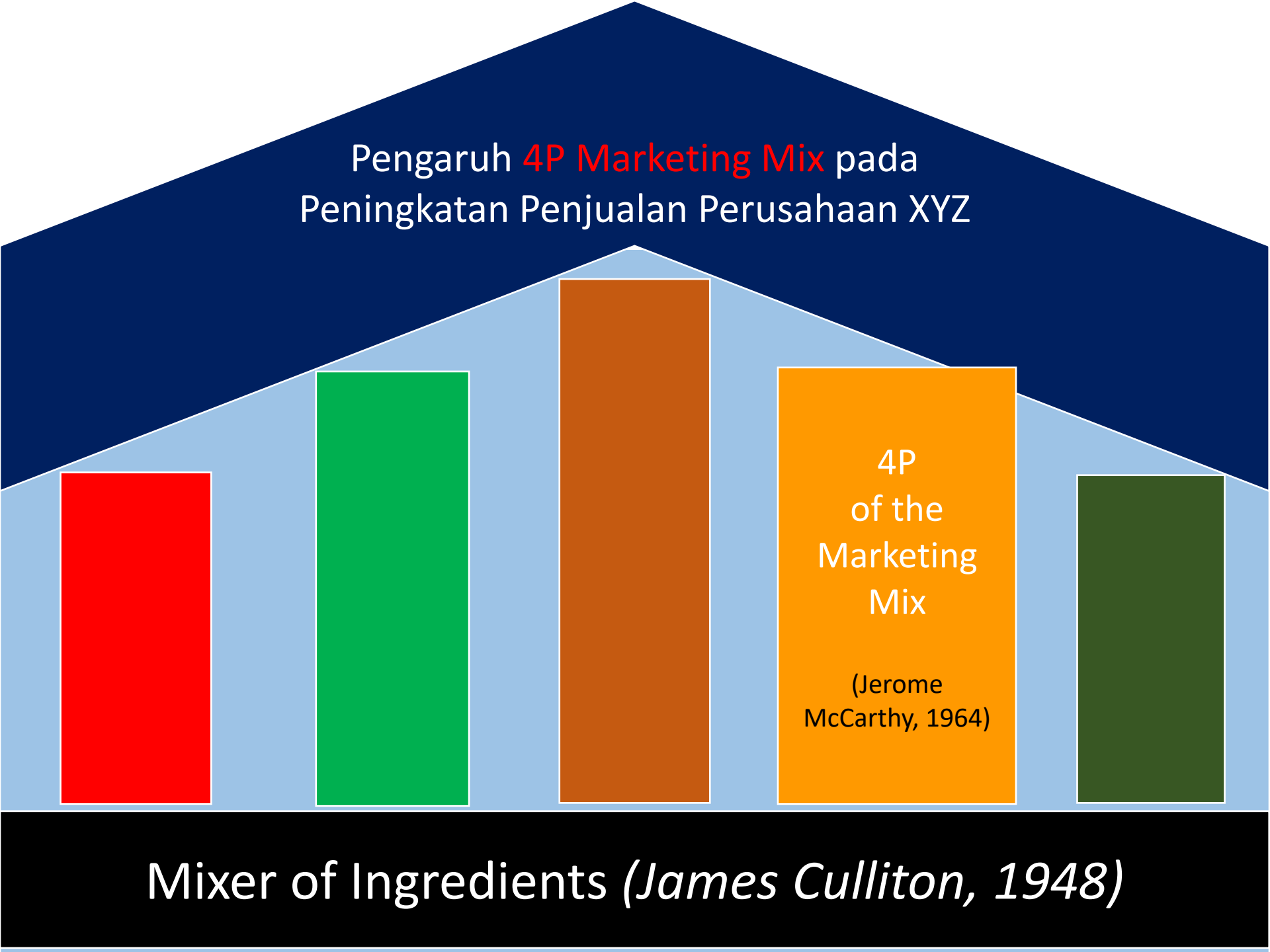


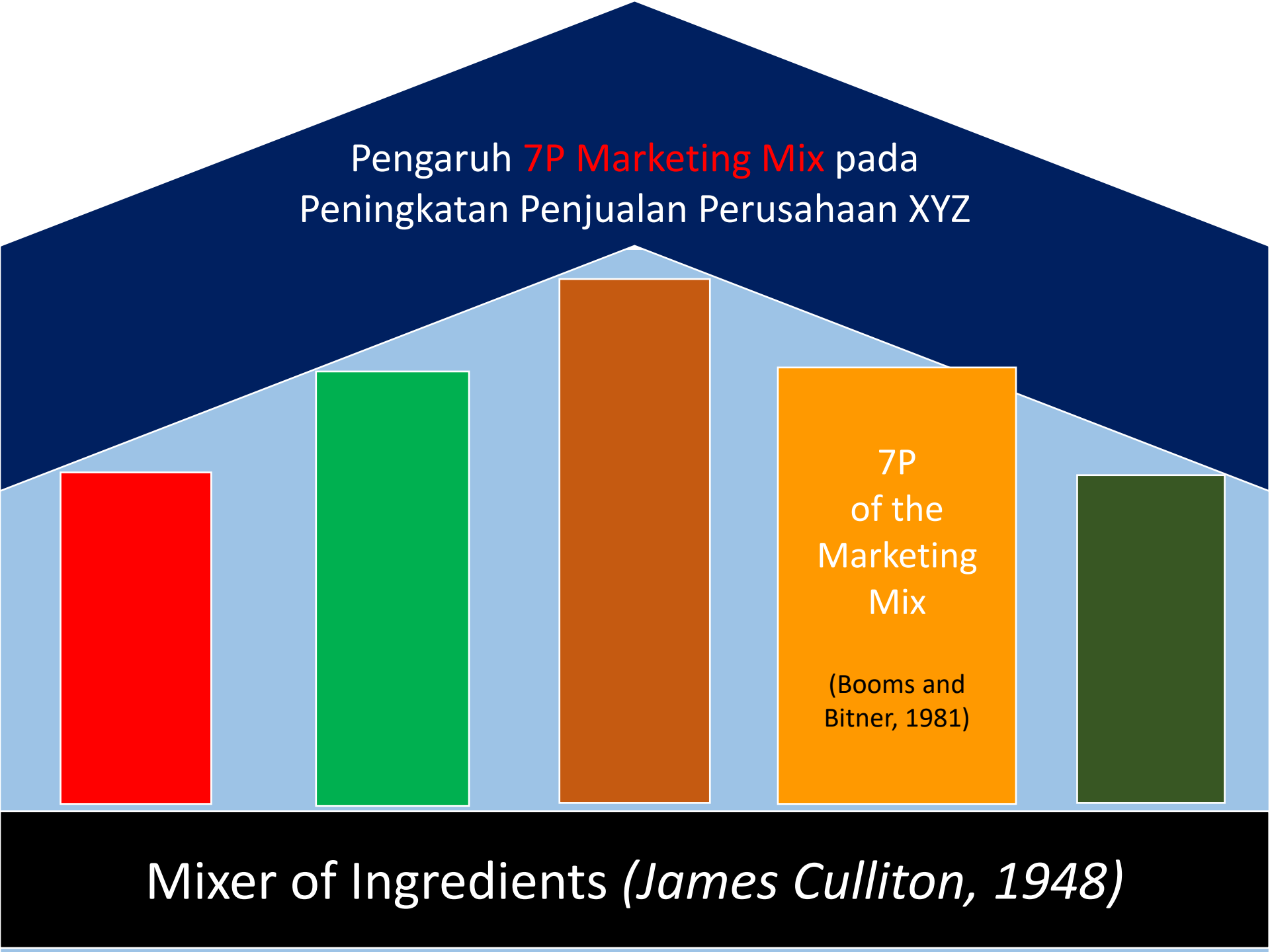
# Marketing Mix Evolution *(Festa et al., 2015)*



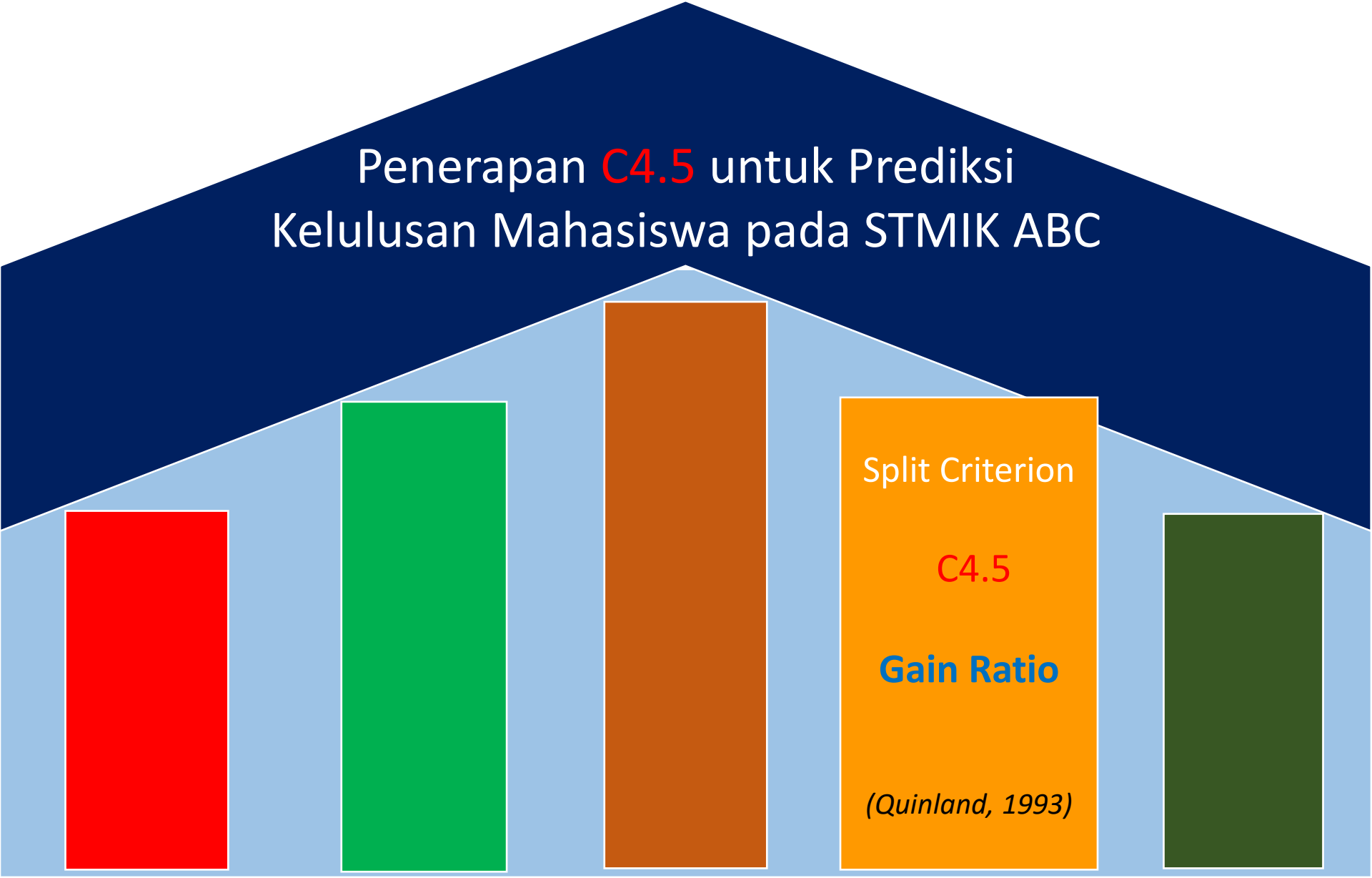

Penelitian Terapan
Penelitian Dasar


Pengaruh 12 Elemen Marketing Mix pada
Peningkatan Penjualan Perusahaan XYZ
12 Elements of the Marketing Mix
(Neil Borden, 1953)
Mixer of Ingredients (James Culliton, 1948)


Pengaruh 4P Marketing Mix pada
Peningkatan Penjualan Perusahaan XYZ
4P
of the
Marketing
Mix
(Jerome
McCarthy, 1964)
Mixer of Ingredients (James Culliton, 1948)

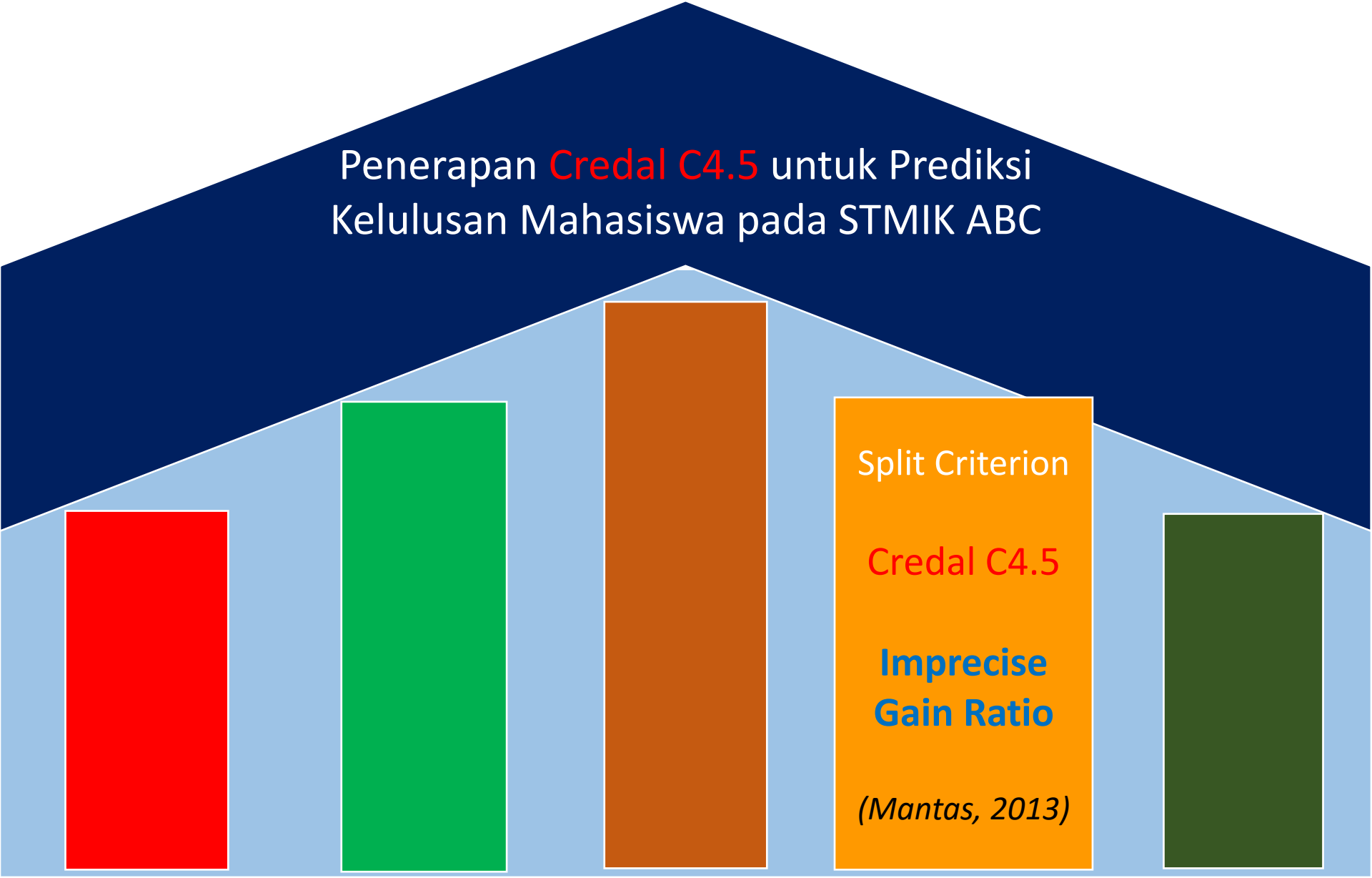

Pengaruh 7P Marketing Mix pada
Peningkatan Penjualan Perusahaan XYZ
7P
of the
Marketing
Mix
(Booms and
Bitner, 1981)
Mixer of Ingredients (James Culliton, 1948)

# Penerapan C4.5 untuk Prediksi Kelulusan Mahasiswa pada STMIK ABC




Teori Gain *(Kullback & Leibler, 1951)*

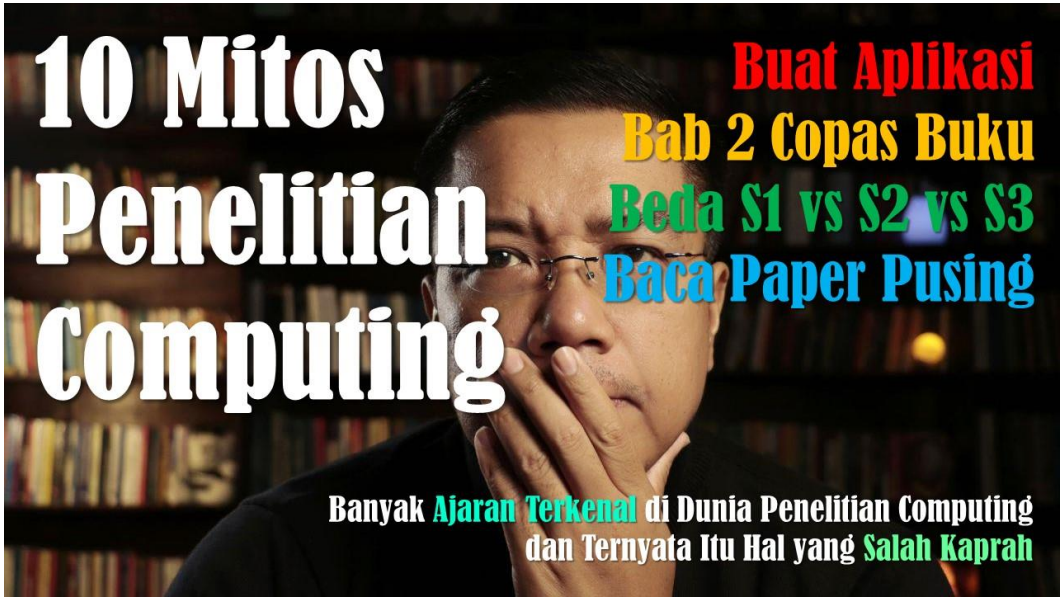
Penerapan Credal C4.5 untuk Prediksi Kelulusan Mahasiswa pada STMIK ABC



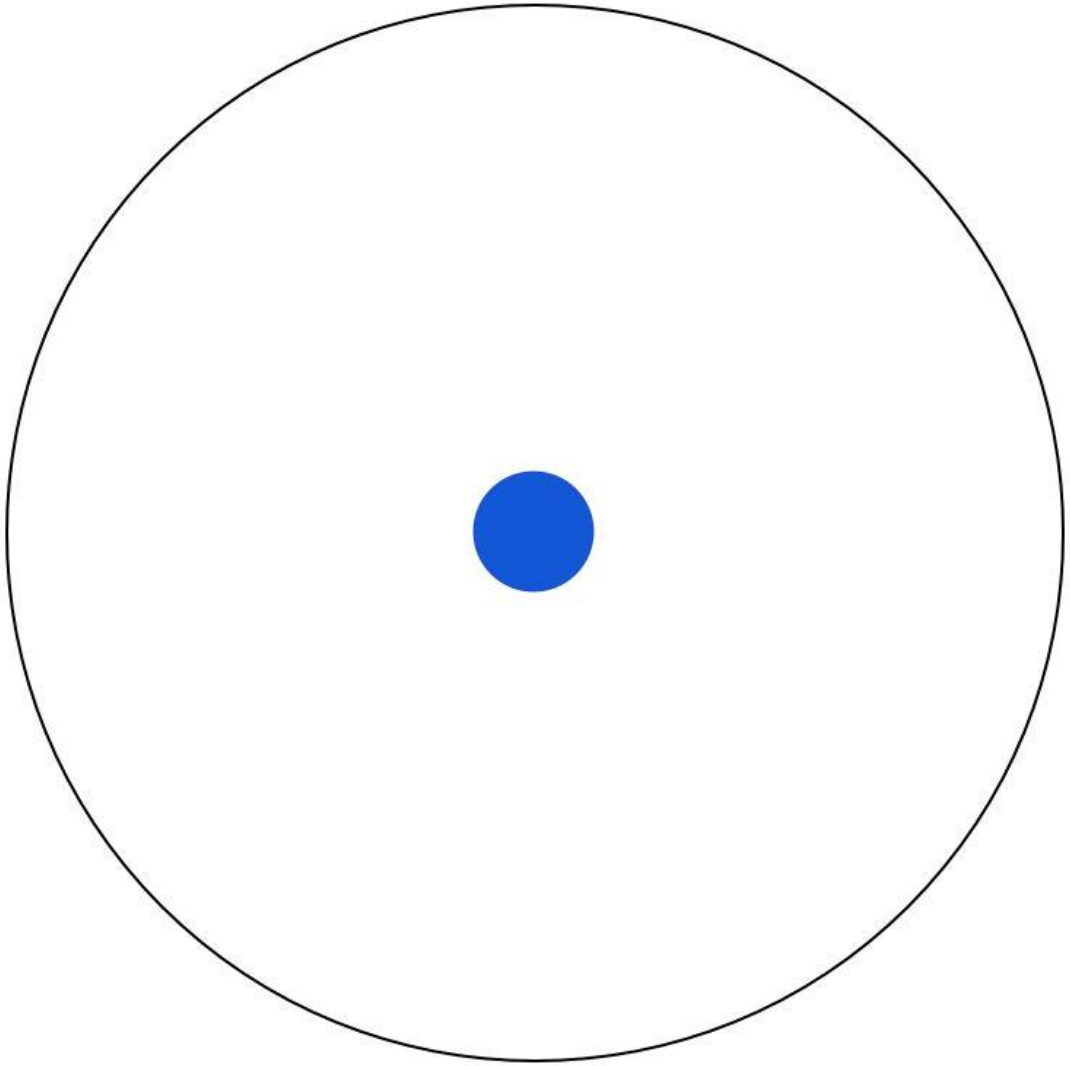

Imprecise Probability Theory *(Walley, 1996)*

# Penelitian yang Berkualitas Tinggi

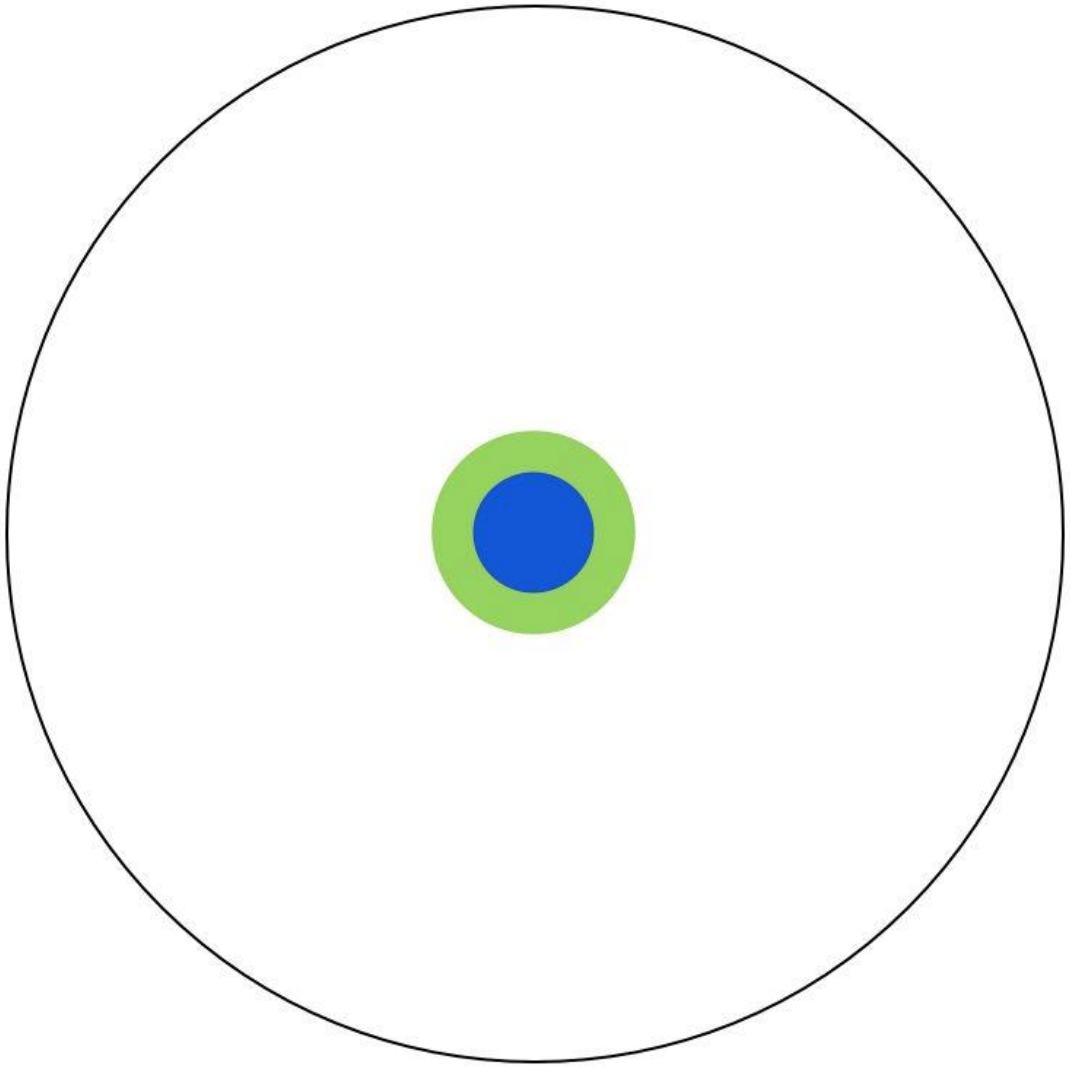
Topik dan skalanya kecil, fokus, dalam, dan membawa pengaruh yang besar ke bidang penelitian kita



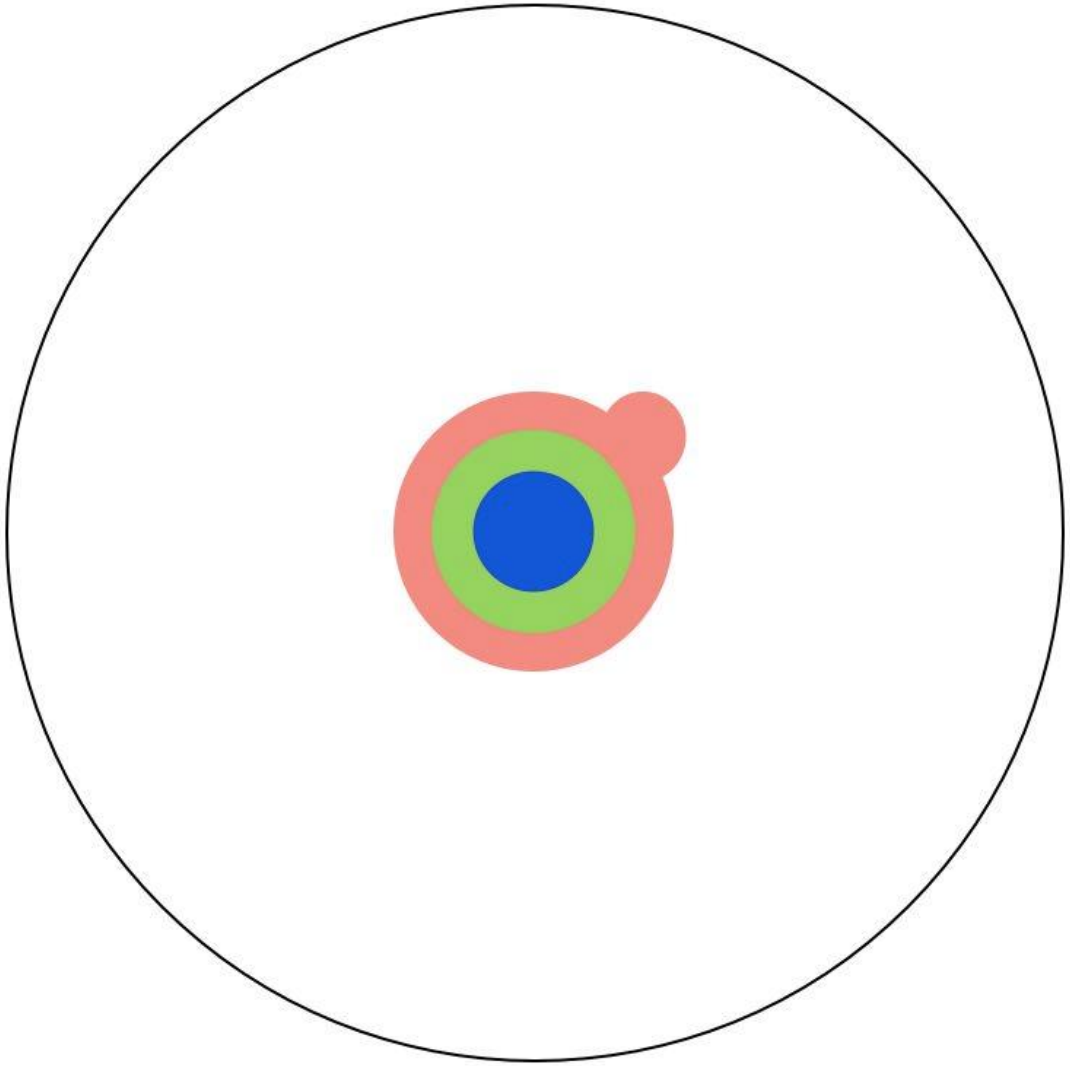
# The Illustrated Guide to a Ph.D (Might, 2010)



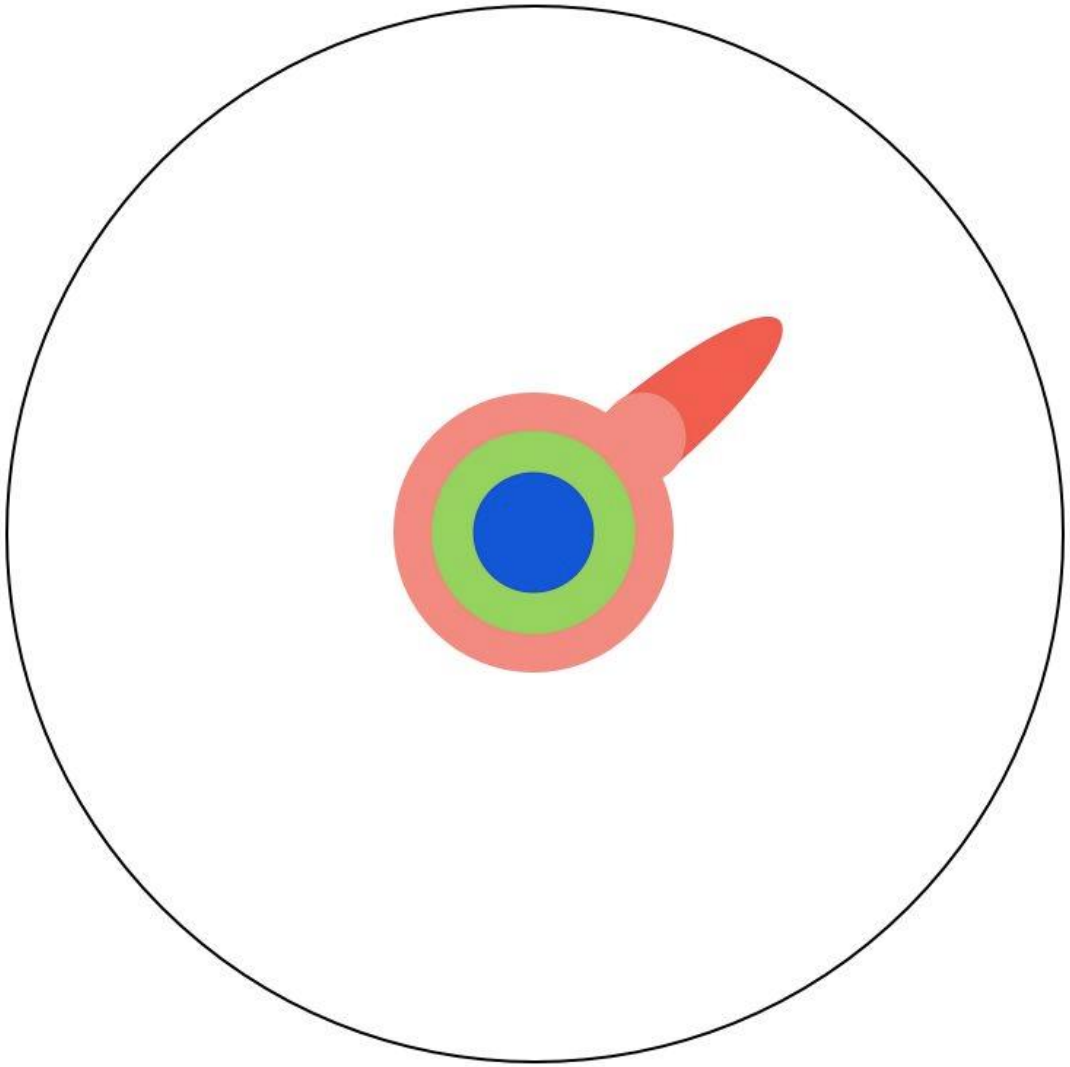
# The Illustrated Guide to a Ph.D (Might, 2010)



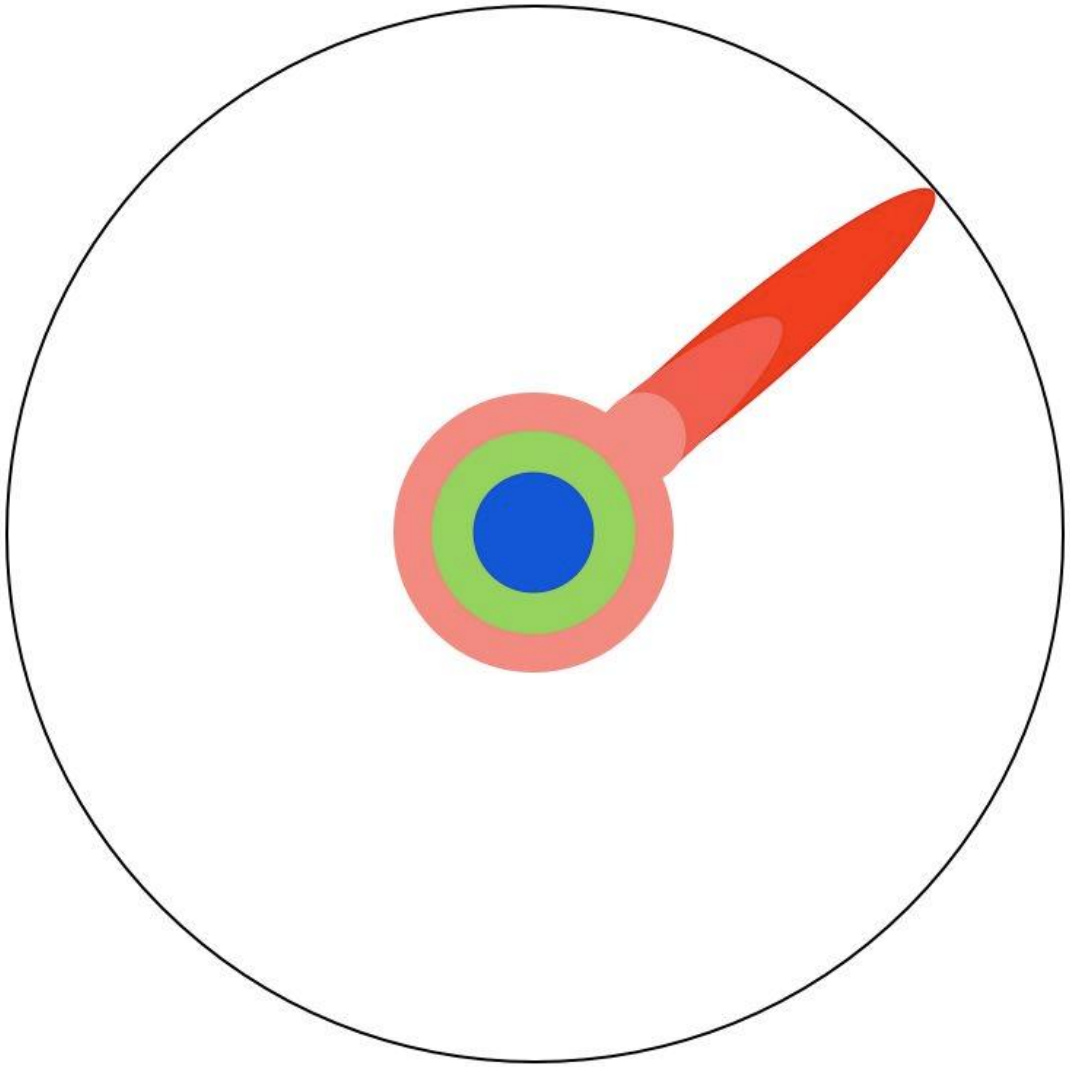
# The Illustrated Guide to a Ph.D (Might, 2010)



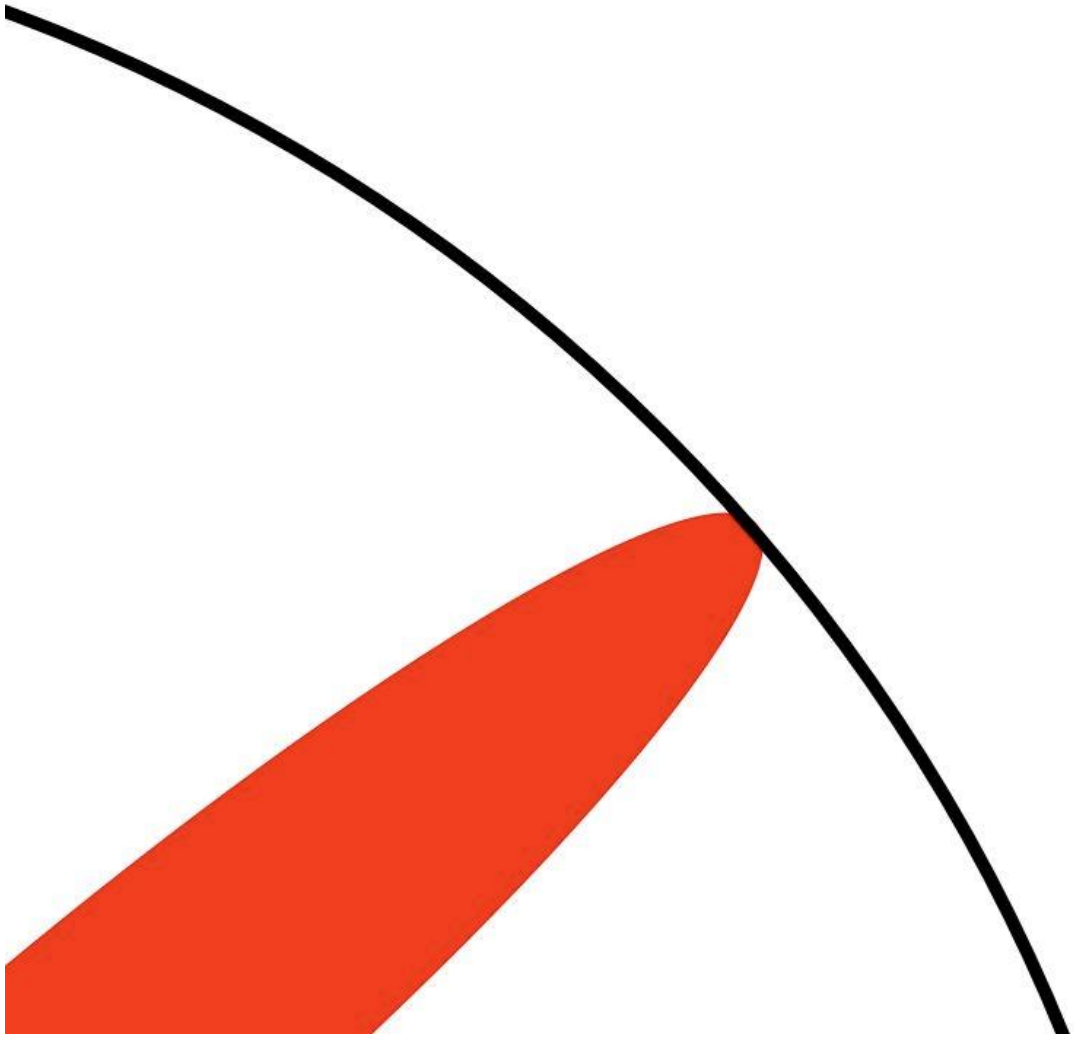
# The Illustrated Guide to a Ph.D (Might, 2010)



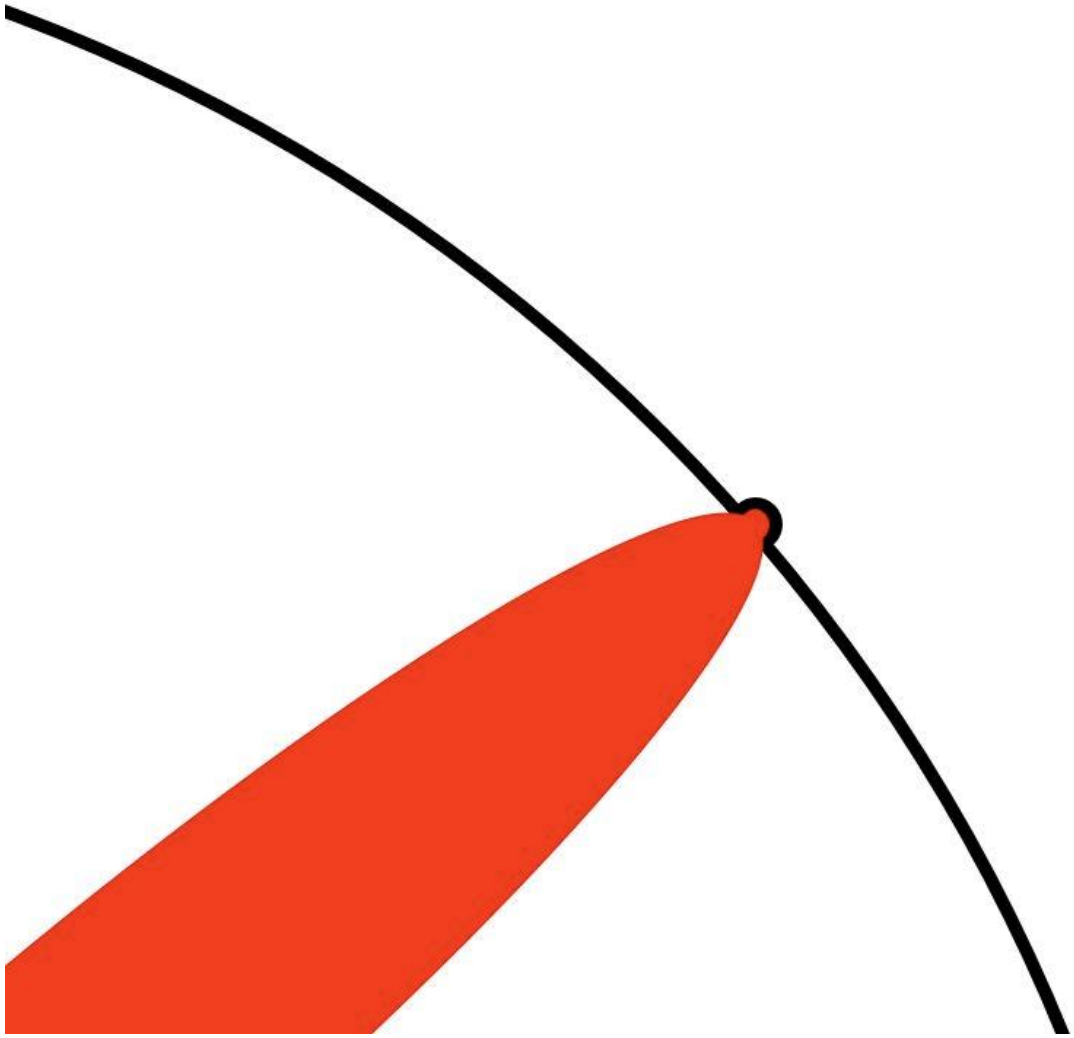
# The Illustrated Guide to a Ph.D (Might, 2010)



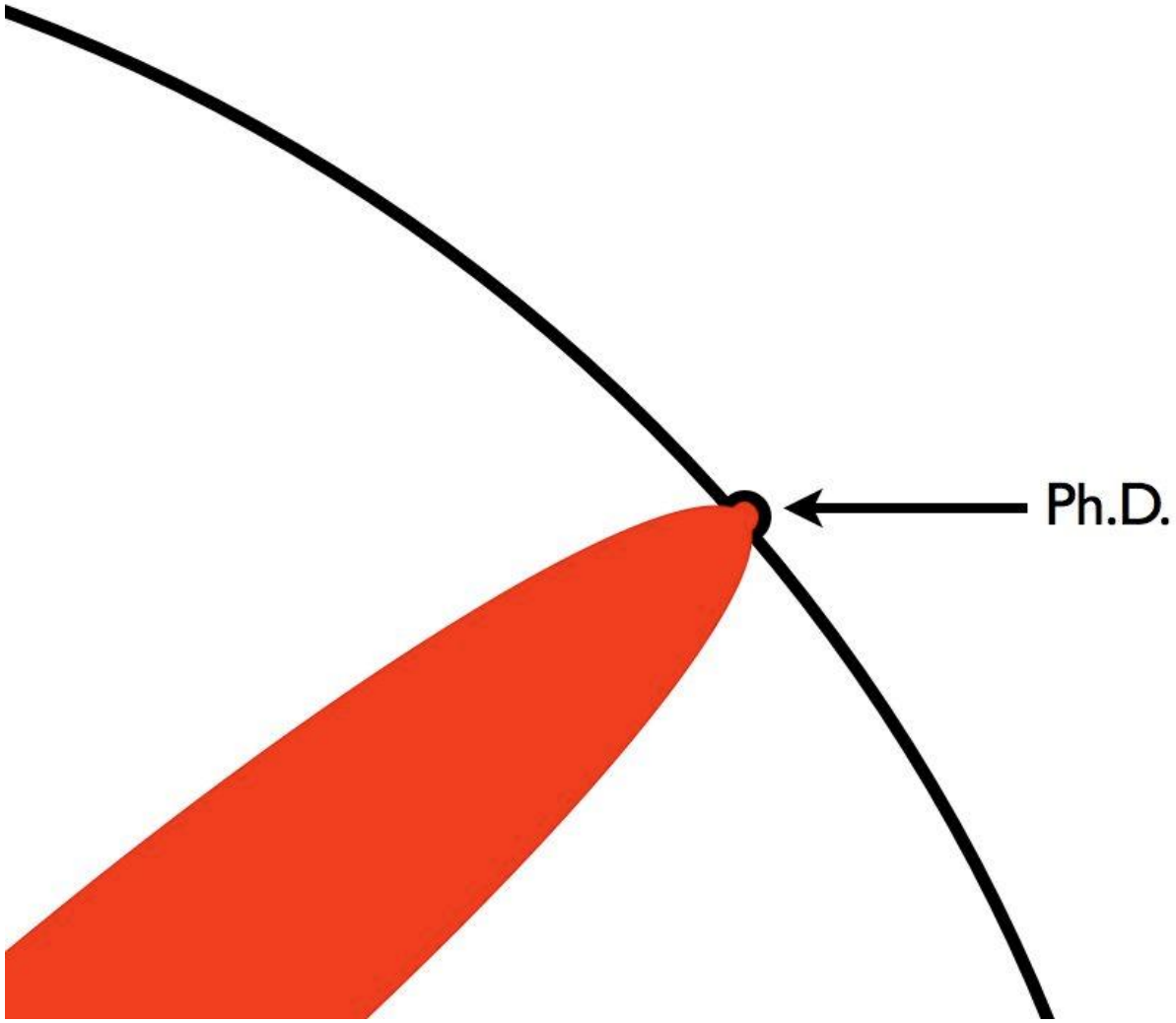
# The Illustrated Guide to a Ph.D (Might, 2010)



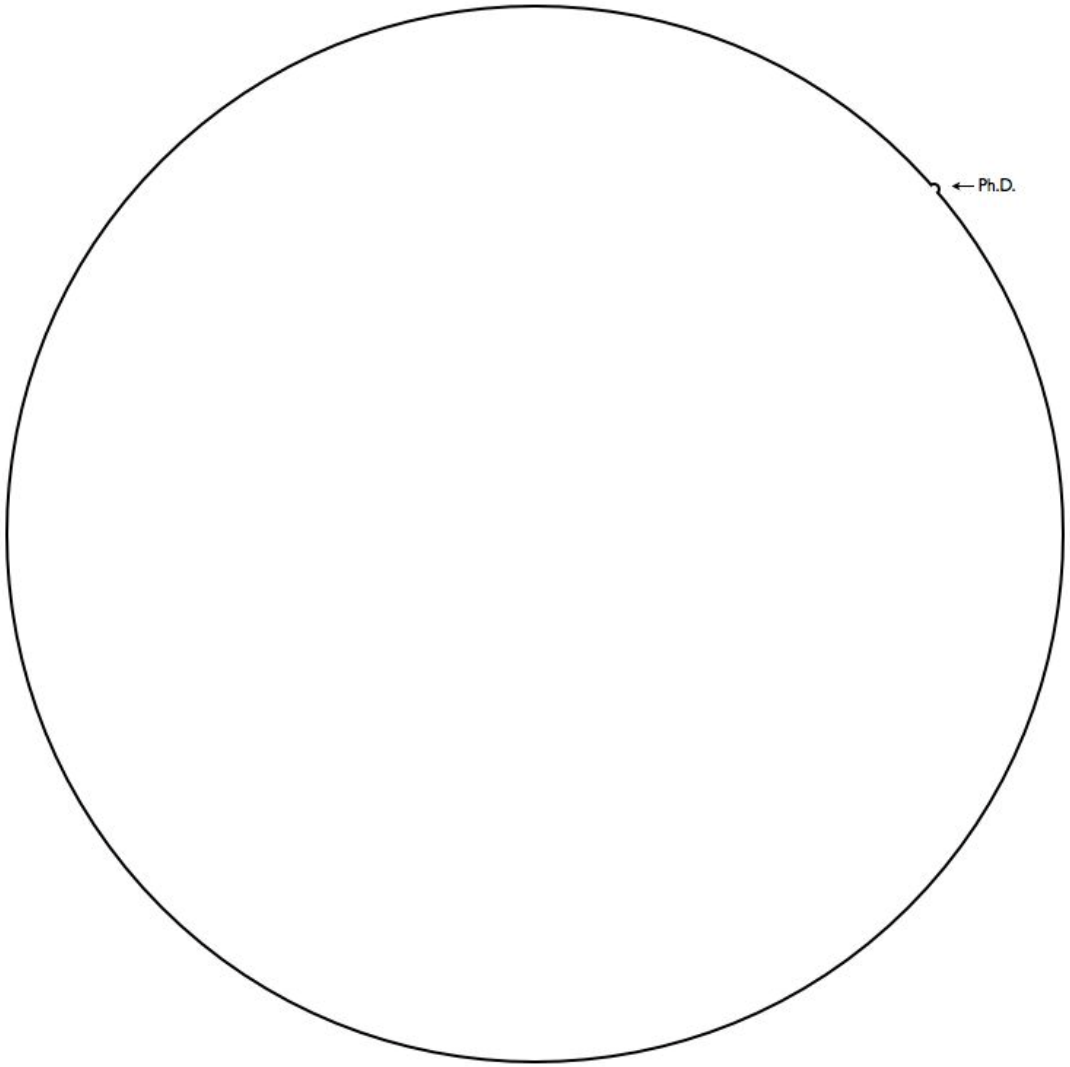
# The Illustrated Guide to a Ph.D (Might, 2010)



# The Illustrated Guide to a Ph.D (Might, 2010)



# The Illustrated Guide to a Ph.D (Might, 2010)



# Akademisi vs Technopreneur

Meja Indah

Meja Kuat

Meja Luas

- Technopreneur?
  1. Jual Produk
  2. Beri Nilai Tambah Produk
  3. Jadikan Aset, Jual Layanan
- Akademisi?
  - Pelajari, Preteli Komponen
  - Ciptakan Meja Baru yang Berbeda dengan 3 Meja Itu

# Apa itu Penelitian?

- Penelitian dilakukan karena ada masalah penelitian, dimana masalah penelitian sendiri muncul karena adanya latar belakang masalah, yang terlahir dari masalah kehidupan
- Penelitian dilakukan secara terencana, sistematis, berulang-ulang dan terukur
- Penelitian harus memiliki orisinalitas (*originality*) dan kebaruan (*novelty*), serta menghasilkan kontribusi yang orisinil pada pengetahuan dalam bentuk menemukan atau merevisi teori, metode, fakta, dan aplikasi

# 1.2 Klasifikasi Penelitian

# Klasifikasi Penelitian

1. **Pendekatan**
   1. Pendekatan Kualitatif
   2. Pendekatan Kuantitatif
2. **Metode**
   1. Metode Penelitian Tindakan
   2. Metode Eksperimen
   3. Metode Studi Kasus
   4. Metode Survei
3. **Jenis**
   1. Dasar vs Terapan
   2. Eksplanatori vs Konfirmatori
   3. Deskripsi vs Eksperimen vs Korelasi
4. **Tingkat**
   1. S1
   2. S2
   3. S3

# 1. Pendekatan

1. Pendekatan Kualitatif:
   - Dari ilmu sosial, konsepnya peningkatan pemahaman terhadap sesuatu, dan bukan membangun penjelasan dari sesuatu *(Berndtsson et al., 2008)*
   - Sifatnya subyektif, berorientasi ke observasi tanpa dikontrol, dan secara umum tidak ada generalisasi *(Blaxter, Hughes, & Tight, 2006)*
   - Dilakukan bidang sistem informasi, dengan metode penelitian seperti "studi kasus" dan "survei", berbasis pola alur induktif

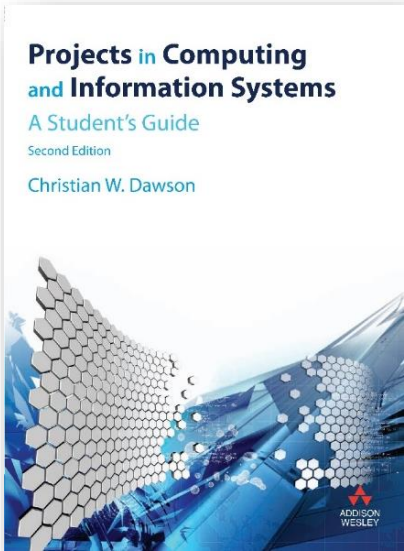

2. Pendekatan Kuantitatif:
   - Dari ilmu alam, konsepnya bagaimana sesuatu dibangun dan bekerja, dan membangun penjelasan dari sesuatu
   - Sifatnya obyektif, berorientasi ke verifikasi, observasi yang dikontrol, dan secara umum ada generalisasi *(Blaxter et al., 2006)*
   - Menggunakan skala numerik, berbasis pola alur deduktif-induktif


*(Berndtsson et al., 2008)*

# 2. Metode

1. Penelitian Tindakan
   - Studi berupa monitoring dan pencatatan penerapan sesuatu oleh peneliti secara hati-hati, yang tujuannya untuk memecahkan masalah dan mengubah situasi *(Herbert, 1990)*
   - Penelitian Tindakan Kelas (PTK) di bidang Pendidikan
2. Eksperimen
   - Investigasi hubungan sebab akibat dengan menggunakan ujicoba yang dikontrol oleh peneliti
   - Melibatkan pengembangan dan evaluasi
   - Penelitian bidang Science dan Teknik
3. Studi Kasus
   - Eksplorasi satu situasi secara mendalam dan hati hati *(Cornford and Smithson, 2006)*
   - Penelitian bidang Sosial, Ekonomi, Politik
4. Survei
   - Pengumpulan data dari populasi yang bisa diukur, dengan cara yang ekonomis *(Saunders et al., 2007)*
   - Melibatkan penggunaan kuesioner dan interview

*(Dawson, 2009)*



# 3. Jenis

Deskripsi

Eksperimen

Korelasi

Kualitatif

Kuantitatif

Eksplanatori

Konfirmatori

Terapan

Dasar

# 4. Tingkat

| Aspek | Skripsi (S1) | Tesis (S2) | Disertasi (S3) |
|---|---|---|---|
| Level Kontribusi | Pengujian Teori | Pengembangan Teori | Penemuan Teori Baru |
| Bentuk Kontribusi | Implementasi dan penerapan | Perbaikan Secara Inkremental dan Terus Menerus | Substansial dan Invention |
| Target Publikasi | Domestic Conference | International Conference | International Journal |

*(Permendikbud No 3 tahun 2020 tentang SNPT)*

# Komparasi Kontribusi Penelitian S1 vs S2 vs S3

- **S1**:
  - Pengaruh 4P Marketing Mix pada Peningkatan Penjualan Perusahaan XYZ
  - *Kontribusi: menguji dan menerapkan teori/hukum/model/metode*
- **S2**:
  - Pengaruh 4P+3C Marketing Mix pada Peningkatan Penjualan Perusahaan XYZ
  - *Kontribusi: mengembangkan dan memperbaiki teori/hukum/model/metode*
- **S3**:
  - Pengaruh ABCD Marketing Mix pada Peningkatan Penjualan Perusahaan
  - *Kontribusi: mengembangan dan menemukan (invention) teori/hukum/model/metode baru yang sifatnya lebih general*

Perbaikan 4P Menjadi 4E

Marketing Mix Baru
Khusus untuk Industri Turisme

Marketing Mix Khusus untuk
Private Labels Brand Equity

Standarisasi Marketing Mix untuk
Generalisasi Lebih Luas

Contents lists available at ScienceDirect

Journal of Business Research

HOSTED BY
ELSEVIER
CrossMark
Available online at www.sciencedirect.com
ScienceDirect
Future Business Journal 3 (2017) 47–69

European Academy of Management and Business Economics
European Journal of Management and Business Economics
www.elsevier.es/ejmbe

Article

Marketing mix effects on

Carmen Abril [a,*], Belén Rodriguez

[a] Facultad de CC Económicas y Empresariales, Universida
[b] Universidad Pontificia de Comillas, C/Alberto Aguilera 2

ARTICLE INFO

*Article history:*
Received 8 June 2016
Accepted 15 September 2016
Available online 25 October 2016

*JEL classification:*
M31
M37

*Keywords:*
Private labels
Store brands
Brand equity
In-store communication
Distribution

**Introduction**

Private label brands, also known as "store brands", were considered low-price, low-q decades ago; currently, however, they repre to manufacturer brands (Kapferer, 2008). than 40% of the market in six European c Manufacturers Association [PLMA], 2015). I refer to brands owned by the retailer or dist

Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
INTERNATIONAL BUSINESS REVIEW
International Busines[illegible]
www.elsevier.com/locate/ibusrev

# Marketing mix standardization: a cross cultural study of four countries

Richard Alan Kustin

*Southern Connecticut State University, New Haven, CT 06515, USA*

**Abstract**

The study researched the possibility of standardizing the marketing mix by investigating the cross-cultural responses from the United States, Brazil, France and India. The study tested the premise of standardization by determining if respondents perceived specific attributes of a common non-durable consumer product the same or differently. The results indicate the opportunity for dynamic marketing standardization remains limited but applicable within specific cultural country markets. Several attribute perceptions between US and foreign respondents are found to be more similar than dissimilar suggesting advantages may exist for a limited implementation of marketing mix standardization as part of a global marketing

# Komparasi Kontribusi Penelitian S1 vs S2 vs S3

- D3/D4:
  - Pengembangan Sistem Informasi Rumah Sakit untuk Rumah Sakit "Suka Sembuh"
  - *Karakter: menguasai skill teknis*
- S1:
  - Sistem Cerdas Berbasis Neural Network untuk Prediksi Harga Saham
  - *Karakter: menguji teori, ada software development*
- S2:
  - Penerapan Algoritma Genetika untuk Pemilihan Arsitektur Jaringan Secara Otomatis pada Neural Network untuk Prediksi Harga Saham
  - *Karakter: mengembangkan teori (perbaikan metode), ada kontribusi ke teori/metode meskipun specific obyek*
- S3:
  - Penerapan Algoritma XYZ untuk Pemilihan Arsitektur Jaringan Secara Otomatis pada Neural Network
  - *Karakter: menemukan teori (invensi metode), ada kontribusi ke teori/metode dengan generalisasi lebih luas*

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Memperbaiki C4.5

## Credal-C4.5: Decision tree based on imprecise probabilities to classify noisy data

Carlos J. Mantas, Joaquín Abellán *

*Department of Computer Science & Artificial Intelligence, University of Granada, ETSI Informática, c/Periodista Daniel Saucedo Aranda s/n, 18071 Granada, Spain*

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Memperbaiki Use Case Points

## Simplifying effort estimation based on Use Case Points ☆

M. Ochodek *, J. Nawrocki, K. Kwarciak

*Poznan University of Technology, Institute of Computing Science, ul. Piotrowo 2, 60-965 Poznań, Poland*



## Genetic Algorithms With Guided and Local Search Strategies for University Course Timetabling

Shengxiang Yang, *Member, IEEE*, and Sadaf Naseem Jat

***Abstract*—The university course timetabling problem (UCTP) is a combinatorial optimization problem, in which a set of events has to be scheduled into time slots and located into suitable rooms. The design of course timetables for academic institutions is a very difficult task because it is an NP-hard problem. This paper investigates genetic algorithms (GAs) with a guided search strategy and local search (LS) techniques for the UCTP. The guided search strategy is used to create offspring into the population based on a data structure that stores information extracted from good individu-**

The research on timetabling problems has a long history of more than 40 years, starting with Gotlieb in 1962 [22]. Researchers have proposed various timetabling approaches by using graph coloring methods, constraint-based methods, population-based approaches (e.g., genetic algorithms (GAs), ant-colony optimization, and memetic algorithms), metaheuristic methods (e.g., tabu search (TS), simulated annealing (SA), and great deluge), variable neighborhood search (VNS), hy-

Memperbaiki Genetic Algorithms

# Parameter Penelitian Yang Berkualitas

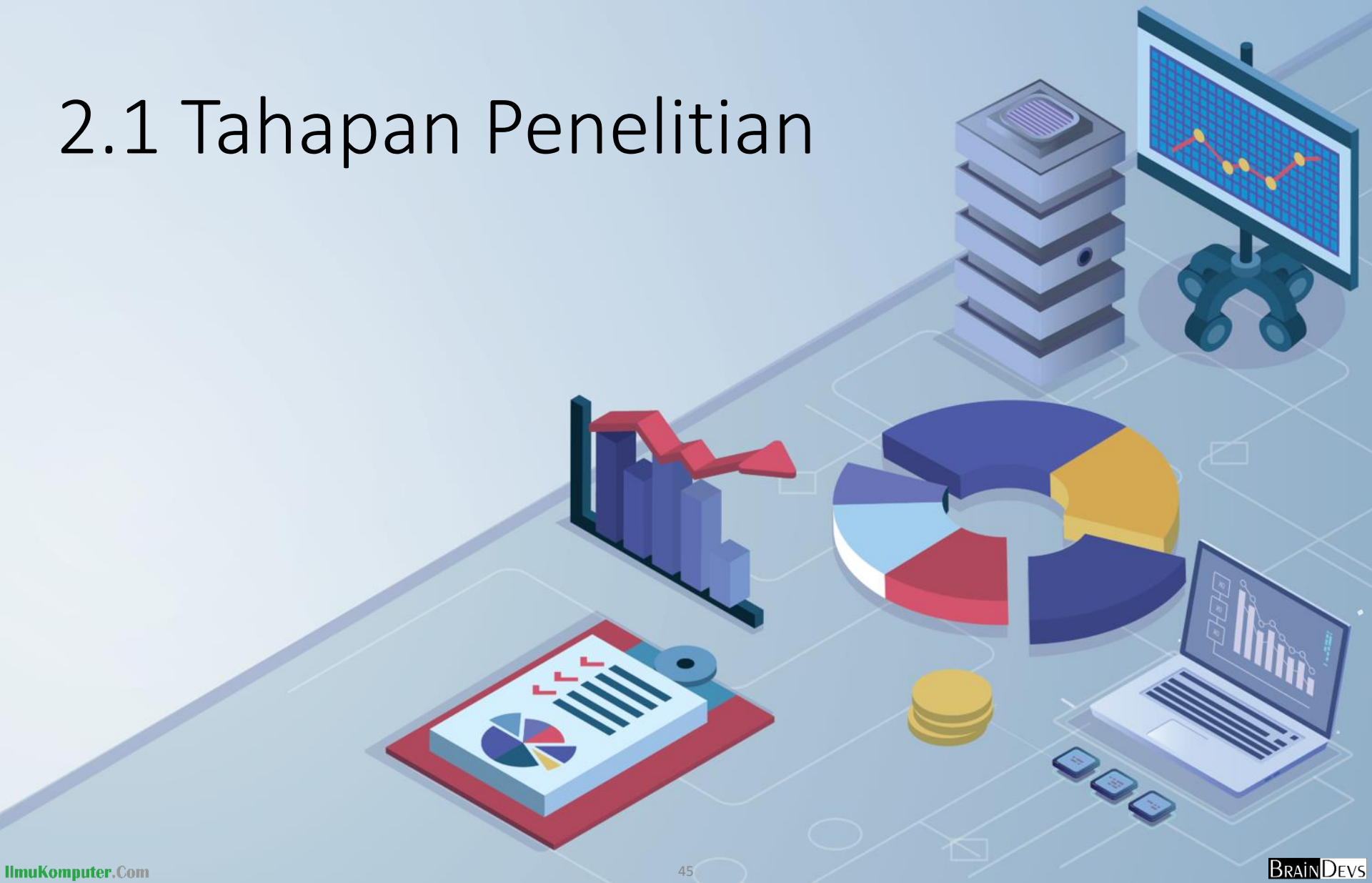
1. Penelitian yang dilakukan secara logis, sistematis, terencana, dan hasil penelitian divalidasi serta terukur *(Supino & Borer, 2012)*
2. Penelitian yang empiris, dilatarbelakangi oleh situasi yang riil, dengan data yang valid dan kongkrit *(Kothari, 2004)*
3. Penelitian yang memiliki kebaruan (*novelty*) yang bisa diwujudkan dalam berbagai bentuk *(Lichtfouse, 2013)*
4. Penelitian yang menghasilkan kontribusi ke pengetahuan yang memiliki orisinalitas yang tinggi *(Sahu, 2013)*
5. Penelitian yang menghasilkan kontribusi ke pengetahuan yang karakternya bisa digeneralisasi untuk obyek yang lain *(Dawson, 2009) (Supino & Borer, 2012)*
6. Penelitian yang bisa direplikasi oleh peneliti lain *(Kothari, 2004) (Runeson et al., 2012)*
7. Penelitian yang mendapatkan sitasi (citation) yang tinggi dari peneliti lain setelah dipublikasi dalam bentuk paper di jurnal ilmiah

# 2. Literature Review

2.1 Tahapan Penelitian

2.2 Literature Review

# 2.1 Tahapan Penelitian



# Tahapan Penelitian Umum

1. Identifikasi Masalah
2. Perumusan Hipotesis
3. Pengujian Hipotesis dan Analisis
4. Kesimpulan

# Tahapan Penelitian Umum vs Skripsi/Tesis

| Tahapan Penelitian | Susunan Tesis |
|---|---|
| 1. Identifikasi **Masalah** | 1. Pendahuluan:<br>- Latar Belakang<br>- Rumusan Masalah<br>- Tujuan Penelitian<br>- Manfaat Penelitian |
| 2. Perumusan **Hipotesis** | 2. Landasan Teori:<br>- Penelitian yang Berhubungan<br>- Landasan Teori<br>- Kerangka Pemikiran |
| 3. **Pengujian Hipotesis** dan Analisis Hasil | 3. Metodologi Penelitian:<br>- Metode Penelitian<br>- Metode Pengumpulan Data<br>- Metode Analisis Data<br>- Metode Pengukuran Penelitian |
| | 4. Analisis Hasil dan Pembahasan |
| 4. **Kesimpulan** | 5. Kesimpulan dan Saran |

# Tahapan Penelitian Komprehensif

## Literature Review

1. Penentuan Bidang Penelitian (***Research Field***)
2. Penentuan Topik Penelitian (***Research Topic***)
3. Penentuan Masalah Penelitian (***Research Problem***)
4. Perangkuman Metode-Metode Yang Ada (***State-of-the-Art Methods***)
5. Penentuan Metode Yang Diusulkan (***Proposed Method***)
6. Evaluasi Metode Yang Diusulkan (***Evaluation***)
7. Penulisan Ilmiah dan Publikasi Hasil Penelitian (***Publications***)

**https://www.site.uottawa.ca/~bochmann/dsrg/how-to-do-good-research/*
**http://romisatriawahono.net/2013/01/23/tahapan-memulai-penelitian-untuk-mahasiswa-galau/*

# 1. Penentuan Bidang Penelitian

- Ingat kembali seluruh mata kuliah yang sudah kita terima di perkuliahan
- Bidang penelitian di disiplin management:

| | |
|---|---|
| Human Resource | Strategic Management |
| Finance | Leadership |
| Marketing | Supply Chain Management |
| Risk Management | Business Process Simulation |
| Project Management | dsb |

- Tentukan berdasarkan passion!

- Contoh: Saya memilih bidang Marketing

# 2. Penentuan Topik Penelitian

- Searching di google, google scholar, ScienceDirect.Com, Springer Link, ACM, IEEE Explore:
  - research trends challenge topics on NAMA BIDANG
- Cara lain menemukan topik adalah dengan menemukan survey/review paper, karena review paper pasti membahas satu topik penelitian
- Contoh:
  - Dari paper-paper survey dan review tentang Marketing, saya tahu trend penelitian di bidang Marketing:
    1. Marketing Mix
    2. Product Development
    3. Branding
    4. Marketing Media
    5. Corporate Social Responsibility
    6. etc
  - Saya mengambil topik penelitian: Marketing Mix

# 3. Penentuan Masalah Penelitian

- Searching di google, google scholar, ScienceDirect.Com:
  - Survey review on NAMA TOPIK
  - Research problem challenge on NAMA TOPIK
- Dari "survey paper" yang ditemukan, kejar sampai dapat semua "technical paper" yang ada di daftar referensinya
- Dari puluhan/ratusan/ribuan paper yang didapat lakukan scanning, pilih paper journal yang terindeks SCOPUS/ISI, 3 tahun terakhir, dan peta kan masalah penelitian yang ada di paper-paper itu
- Gunakan Mendeley untuk mempermudah pekerjaan kita
- Pilih satu atau dua masalah penelitian yang kita anggap menarik dan menantang, dan jadikan itu masalah penelitian kita

# 4. Perangkuman Metode Yang Ada

- Pahami semua paper penelitian yang tujuannya memecahkan masalah yang sama dengan yang kita pilih
- Pahami teori/hukum/metode/algoritma terkini yang mereka gunakan untuk memecahkan masalah penelitian mereka
  - Ini yang disebut dengan state-of-the-art method
- Dalam bidang computing, metode biasanya berupa algoritma yang secara sistematis, logis dan matematis menyelesaikan masalah

# State-of-the-Art Frameworks in Software Defect Prediction

**Menzies** Framework

(Menzies et al. 2007)

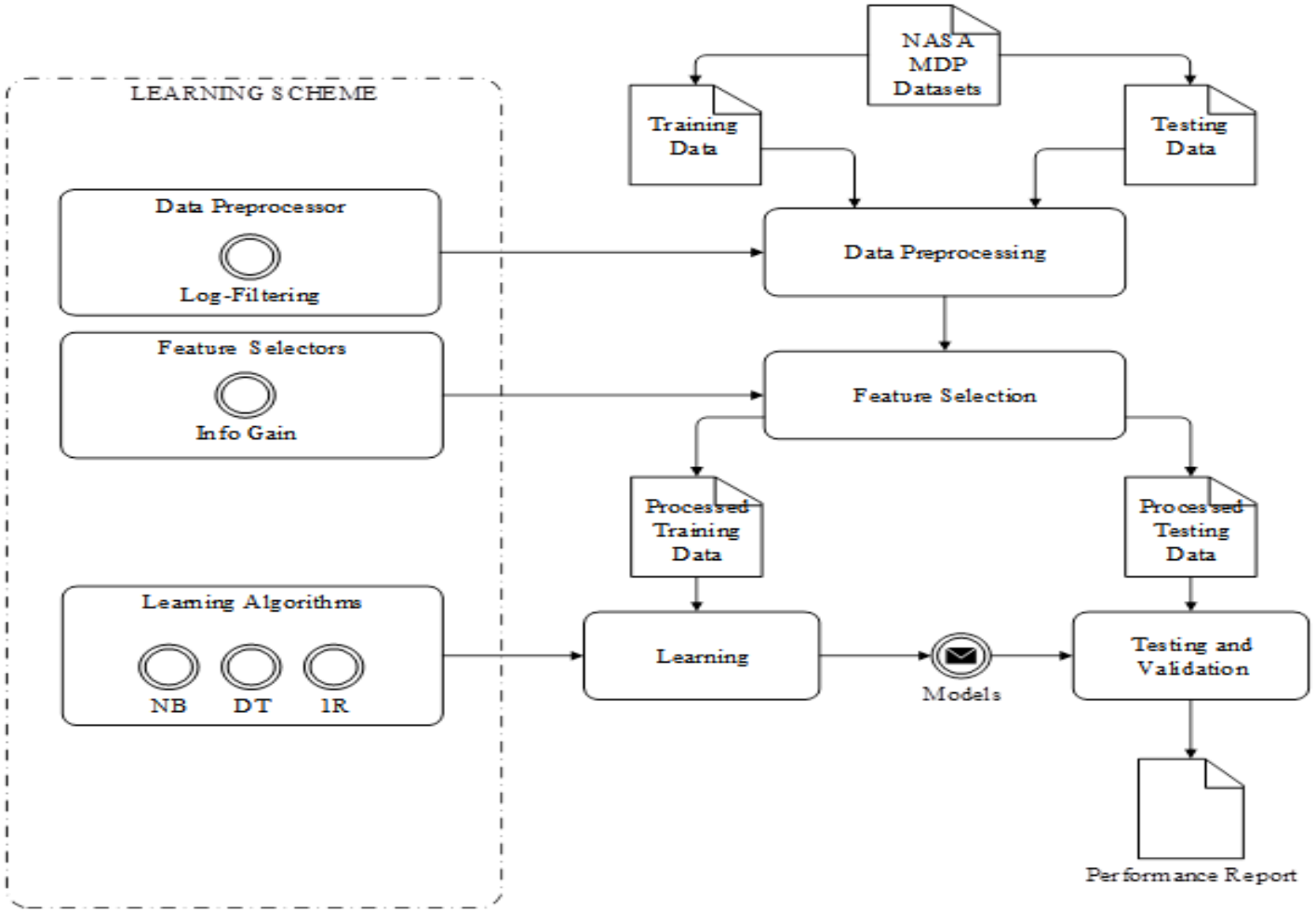
**Lessmann** Framework

(Lessmann et al. 2008)

**Song** Framework

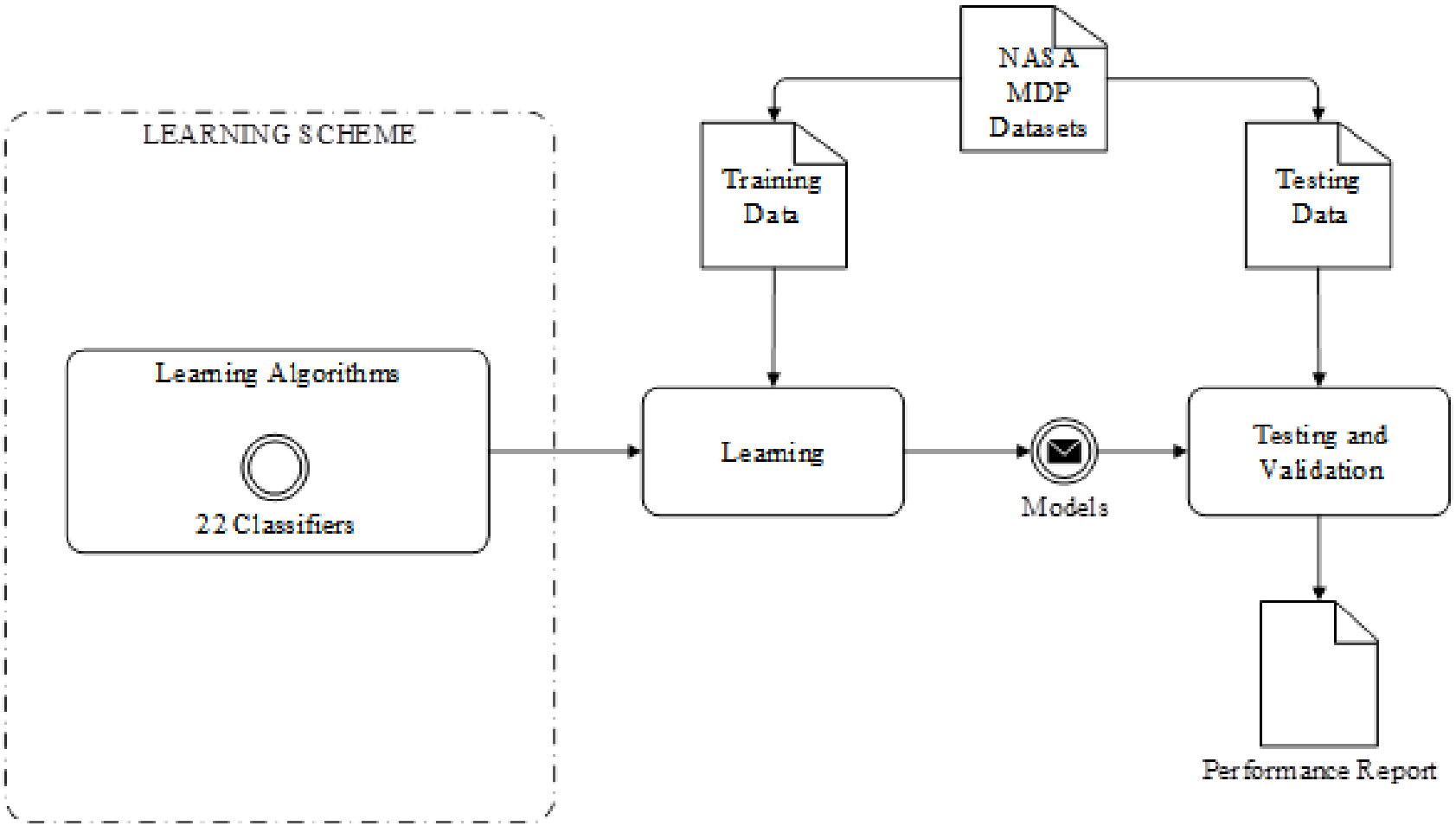
(Song et al. 2011)

# Menzies Framework (Menzies et al. 2007)



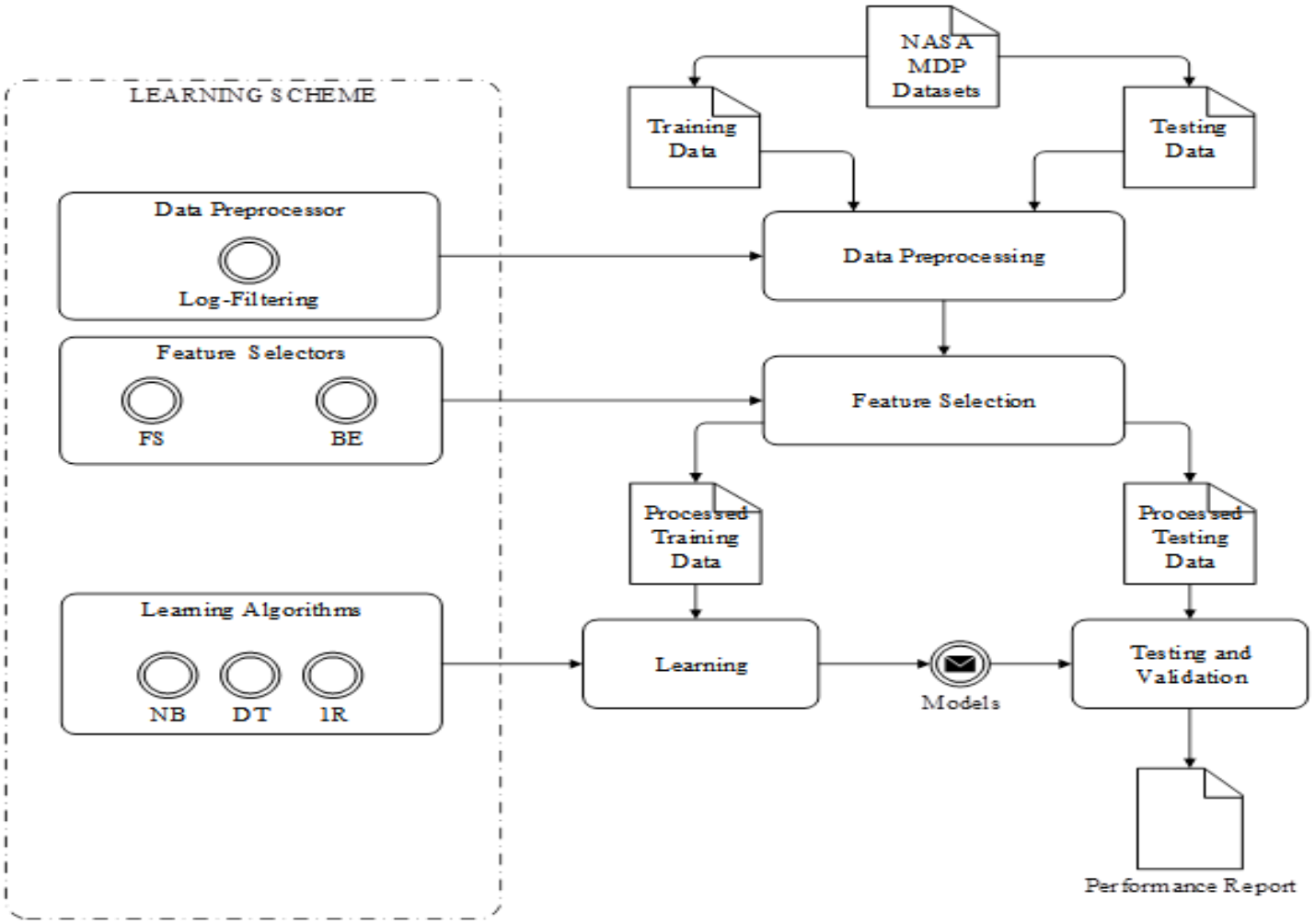

| Framework | Dataset | Data Preprocessor | Feature Selectors | Meta-learning | Classifiers | Parameter Selectors | Validation Methods | Evaluation Methods |
|---|---|---|---|---|---|---|---|---|
| (Menzies et al. 2007) | NASA MDP | **Log Filtering** | **Info Gain** | - | **3 algorithms (DT, 1R, NB)** | - | 10-Fold X Validation | ROC Curve (AUC) |

# Lessmann Framework (Lessmann et al. 2008)




| Framework | Dataset | Data Preprocessor | Feature Selectors | Meta-learning | Classifiers | Parameter Selectors | Validation Methods | Evaluation Methods |
|---|---|---|---|---|---|---|---|---|
| (Lessman et al. 2008) | NASA MDP | - | - | - | **22 algorithms** | - | 10-Fold X Validation | ROC Curve (AUC) |

# Song Framework (Song et al. 2011)

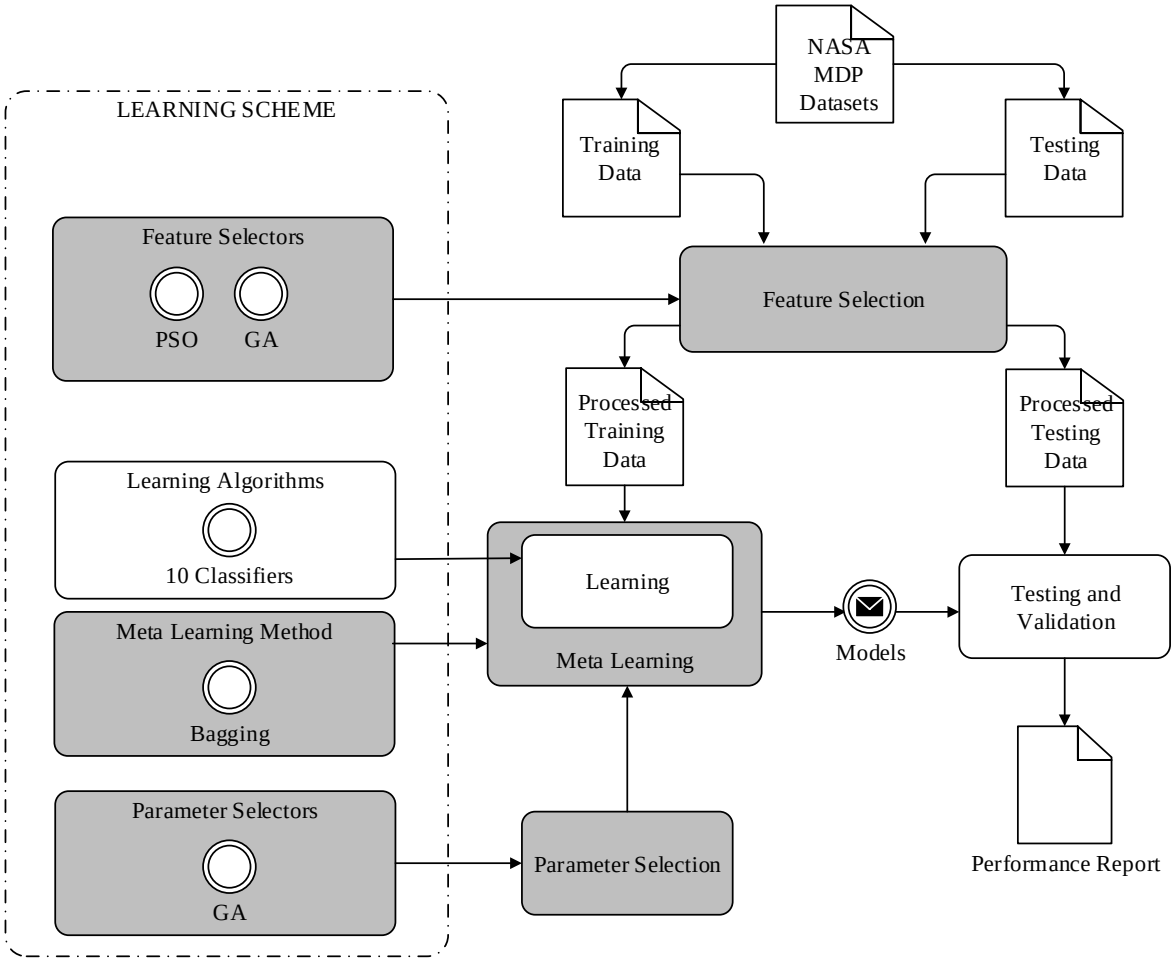



| Framework | Dataset | Data Preprocessor | Feature Selectors | Meta-learning | Classifiers | Parameter Selectors | Validation Methods | Evaluation Methods |
|---|---|---|---|---|---|---|---|---|
| (Song et al. 2011) | NASA MDP | **Log Filtering** | **FS, BE** | - | **3 algorithms (DT, 1R, NB)** | - | 10-Fold X Validation | ROC Curve (AUC) |

# 5. Penentuan Metode Yang Diusulkan

- Kita harus membangun dan mengusulkan suatu metode (*proposed method*), yg lebih baik bila dibandingkan dengan metode-metode yang ada saat ini
- Keunggulan metode yang kita usulkan harus dilandasi (*reference*), dibuktikan secara matematis dan empiris lewat hasil eksperimen dan perbandingan dengan metode yang ada
- Metode yang kita usulkan itu bisa saja dari *state-of-the-art methods*, kita kemudian "menambahkan" sesuatu (algoritma, koefisien, formula, dsb), yang akhirnya ketika kita bandingkan dengan metode original, metode kita lebih baik (lebih cepat, lebih akurat, lebih konsisten, dsb).
- "Penambahan" yang kita lakukan dan akhirnya membuat pemecahan masalah menjadi lebih baik itulah yang disebut dengan kontribusi ke pengetahuan (*contribution to knowledge*) (Dawson, 2009)

# Proposed Framework




| Framework | Dataset | Data Preprocessor | Feature Selectors | Meta-Learning | Classifiers | Parameter Selectors | Validation Methods | Evaluation Methods |
|---|---|---|---|---|---|---|---|---|
| (Menzies et al. 2007) | NASA MDP | Log Filtering | Info Gain | | 3 algorithm (DT, 1R, NB) | - | 10-Fold X Validation | ROC Curve (AUC) |
| (Lessman et al. 2008) | NASA MDP | - | - | | 22 algorithm | - | 10-Fold X Validation | ROC Curve (AUC) |
| (Song et al. 2011) | NASA MDP | Log Filtering | FS, BE | | 3 algorithm (DT, 1R, NB) | - | 10-Fold X Validation | ROC Curve (AUC) |
| Proposed Framework | NASA MDP | - | **PSO, GA** | **Bagging** | **10 algorithms** | **GA** | 10-Fold X Validation | ROC Curve (AUC) |

# 6. Evaluasi Metode Yang Diusulkan

- Metode yang diusulkan harus divalidasi dan dievaluasi dengan metode pengukuran standard dan disepakati para peneliti di bidang penelitian yang kita lakukan
  - Ujungnya supaya bisa ditarik kesimpulan yang menghasilkan finding (contribution to knowledge)
- Di dunia computing, pengukuran metode disesuaikan dengan masalah dan tujuan penelitian:
  - Masalahnya rendahnya akurasi → pengukurannya akurasi
  - Masalah rendahnya efisiensi → pengukurannya waktu

# 7. Penulisan Ilmiah dan Publikasi Hasil Penelitian

- Lakukan pendataan journal-journal yang ada di bidang kita, urutkan berdasarkan rangking SJR atau JIF
- Pilih target journal untuk tempat publikasi hasil penelitian kita
- Publikasikan hasil penelitian ke journal yang sesuai dengan kualitas kontribusi penelitian yang kita lakukan
- A paper is an organized description of hypotheses, data and conclusions, intended to instruct the reader. If your research does not generate papers, it might just as well not have been done (Whitesides 2004)

| No | Journal Publications | SJR | Q Category |
|---|---|---|---|
| 1 | IEEE Transactions on Software Engineering | 3.39 | Q1 in Software |
| 2 | Information Sciences | 2.96 | Q1 in Information Systems |
| 3 | IEEE Transactions on Systems, Man, and Cybernetics | 2.76 | Q1 in Artificial Intelligence |
| 4 | IEEE Transactions on Knowledge and Data Engineering | 2.68 | Q1 in Information Systems |
| 5 | Empirical Software Engineering | 2.32 | Q1 in Software |
| 6 | Information and Software Technology | 1.95 | Q1 in Information Systems |
| 7 | Automated Software Engineering | 1.78 | Q1 in Software |
| 8 | IEEE Transactions on Reliability | 1.43 | Q1 in Software |
| 9 | Expert Systems with Applications | 1.36 | Q2 in Computer Science |
| 10 | Journal of Systems and Software | 1.09 | Q2 in Software |
| 11 | Software Quality Journal | 0.83 | Q2 in Software |
| 12 | IET Software | 0.55 | Q2 in Software |
| 13 | Advanced Science Letters | 0.24 | Q3 in Computer Science |
| 14 | Journal of Software | 0.23 | Q3 in Software |
| 15 | International Journal of Software Engineering and Its Application | 0.14 | Q4 in Software |

# 2.2 Literature Review

# Tahapan Penelitian

## Literature Review

1. Penentuan Bidang Penelitian (***Research Field***)
2. Penentuan Topik Penelitian (***Research Topic***)
3. Penentuan Masalah Penelitian (***Research Problem***)
4. Perangkuman Metode-Metode Yang Ada (***State-of-the-Art Methods***)
5. Penentuan Metode Yang Diusulkan (***Proposed Method***)
6. Evaluasi Metode Yang Diusulkan (***Evaluation***)
7. Penulisan Ilmiah dan Publikasi Hasil Penelitian (***Publications***)

**https://www.site.uottawa.ca/~bochmann/dsrg/how-to-do-good-research/*
**http://romisatriawahono.net/2013/01/23/tahapan-memulai-penelitian-untuk-mahasiswa-galau/*

# Manfaat Mereview Literatur

- Memperdalam pengetahuan tentang bidang dan topik yang diteliti (*Textbooks*)
- Memperdalam pengetahuan tentang topik lebih detail yang diteliti (*Survey Paper*)
- Mengetahui hasil penelitian yang berhubungan dan yang sudah pernah dilaksanakan (Related Research) (*Technical Paper*)
- Mengetahui perkembangan ilmu pada bidang yang kita pilih (state-of-the-art) (*Technical atau Survey Paper*)
- Memperjelas masalah penelitian (*Technical Paper*)

# Jenis Literatur Ilmiah

1. **Paper dari Journal** *
2. Paper dari Book Chapter
3. Paper dari Conference (Proceedings)
4. Thesis dan Disertasi
5. Report (Laporan) dari Organisasi yang Terpercaya
6. Buku Textbook


* *Prioritaskan paper journal terindeks WoS (Clarivate) or SCOPUS*
*Cek dengan http://scimagojr.com*

SJR SCImago Journal & Country Rank

EST MODUS IN REBUS
Horatio (Satire 1,1,106)

- Home
- Journal Rankings
- Journal Search
- Country Rankings
- Country Search
- Compare
- Map Generator
- Help
- About Us

## Journal Search

Search query

Journal of Systems and Software in Journal Title Search

Exact phrase

**Please, select journal:**

1. Journal of Systems and Software. United States.

## Related product

SJR SCImago Journal & Country Rank

EST MODUS IN REBUS
Horatio (Satire 1,1,106)

Home
Journal Rankings
Journal Search
Country Rankings
Country Search
Compare
Map Generator
Help
About Us

# Journal Search

Search query

in Journal Title Search

Exact phrase

# Journal of Systems and Software

**Country:** United States

**Subject Area:** Computer Science

**Subject Category:**

| Category | Quartile (Q1 means highest values and Q4 lowest values) | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
| Hardware and Architecture | Q2 | Q2 | Q1 | Q2 | Q2 | Q2 | Q1 | Q2 | Q2 | Q2 | Q1 | Q2 | Q1 | Q1 | Q1 |
| Information Systems | Q3 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 |
| Software | Q3 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 | Q2 |

**Publisher:** Elsevier Inc.. **Publication type:** Journals. **ISSN:** 01641212

**Coverage:** 1979-2014

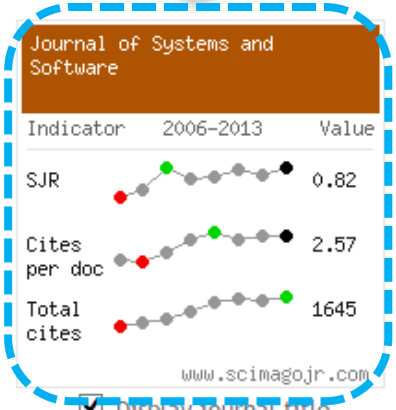
**H Index:** 60

**Scope:**

The Journal of Systems and Software publishes papers covering all aspects of programming methodology, software engineering, and related hardware-software-systems issues. [...]

Show full scope

Show this information in your own website




Display journal title

# Sumber Pencarian Literatur

## GRATIS

**Journal**

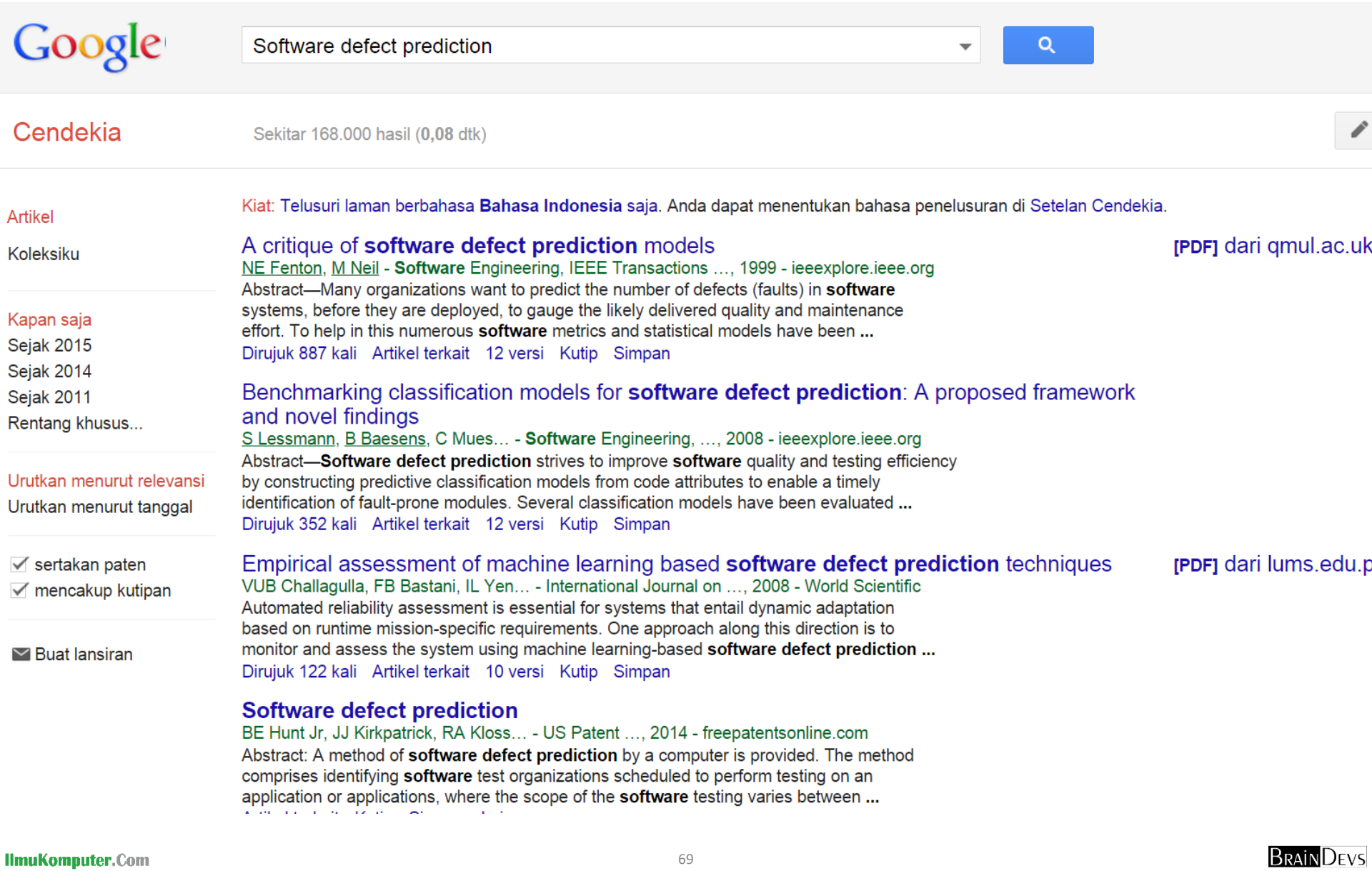
- http://scholar.google.com
- http://citeseer.ist.psu.edu

## BERBAYAR

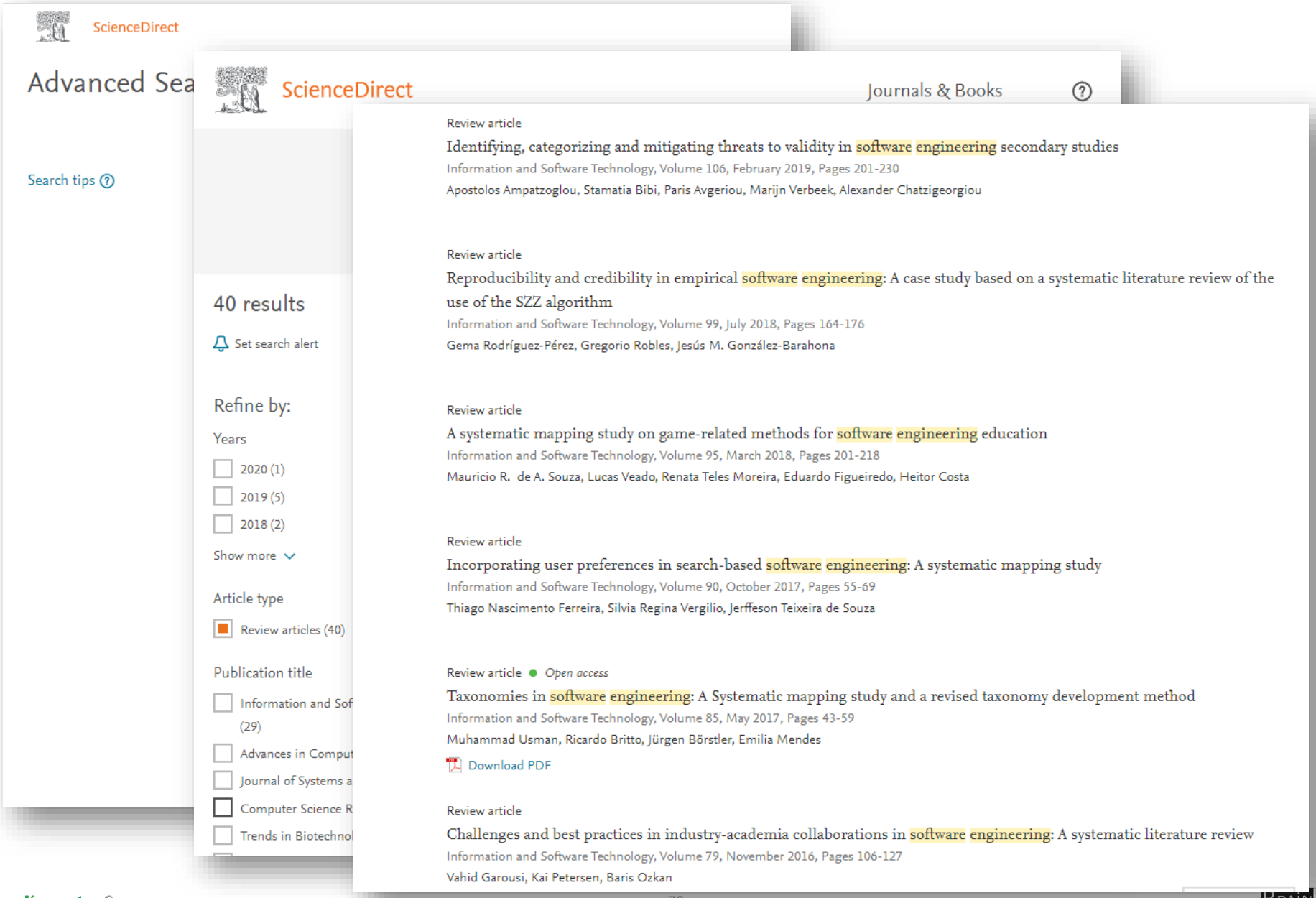
**Journal**

- http://sciencedirect.com
- http://www.ebscohost.com
- http://link.springer.com
- http://ieeexplore.ieee.org
- http://dl.acm.org

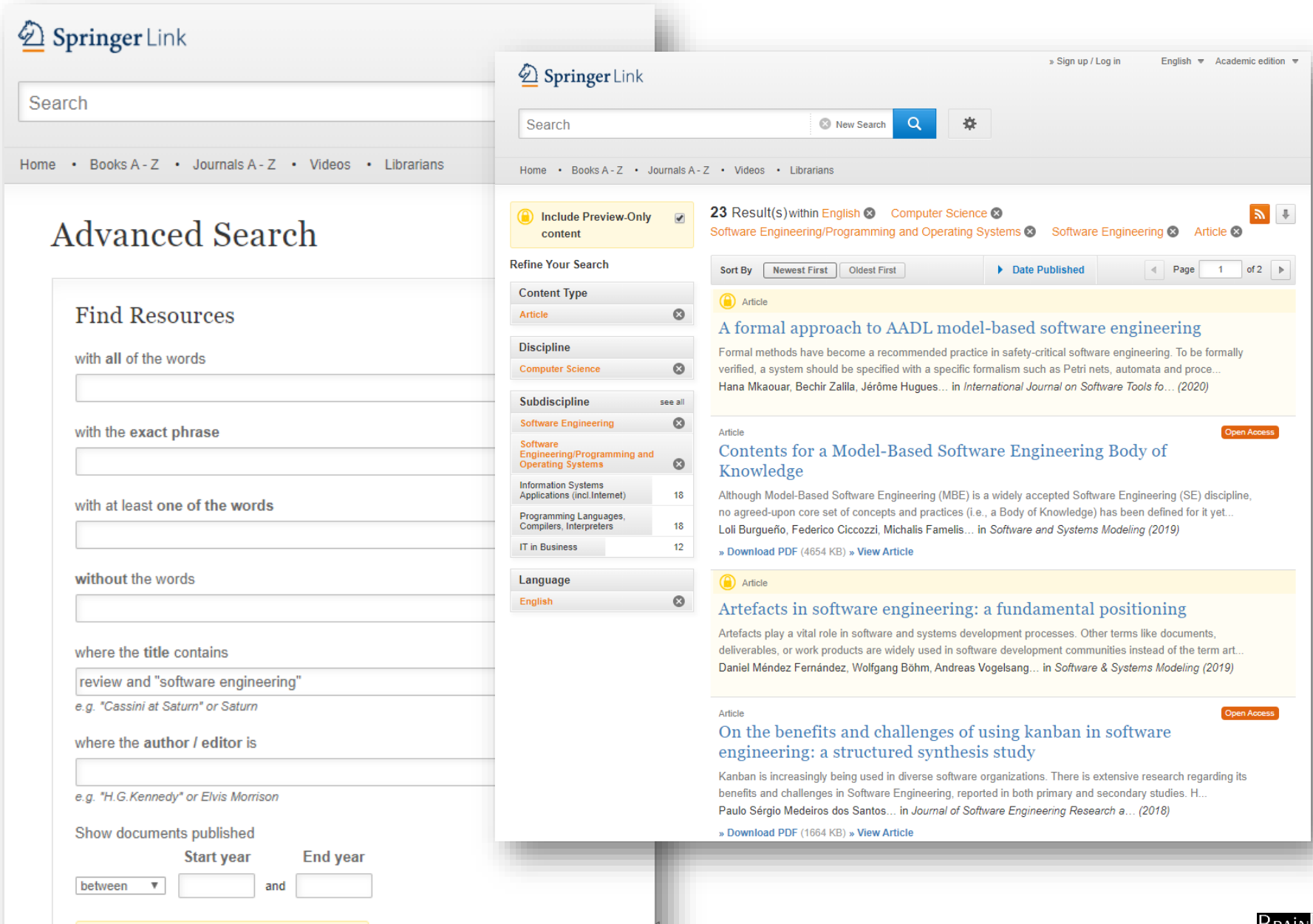
# scholar.google.com



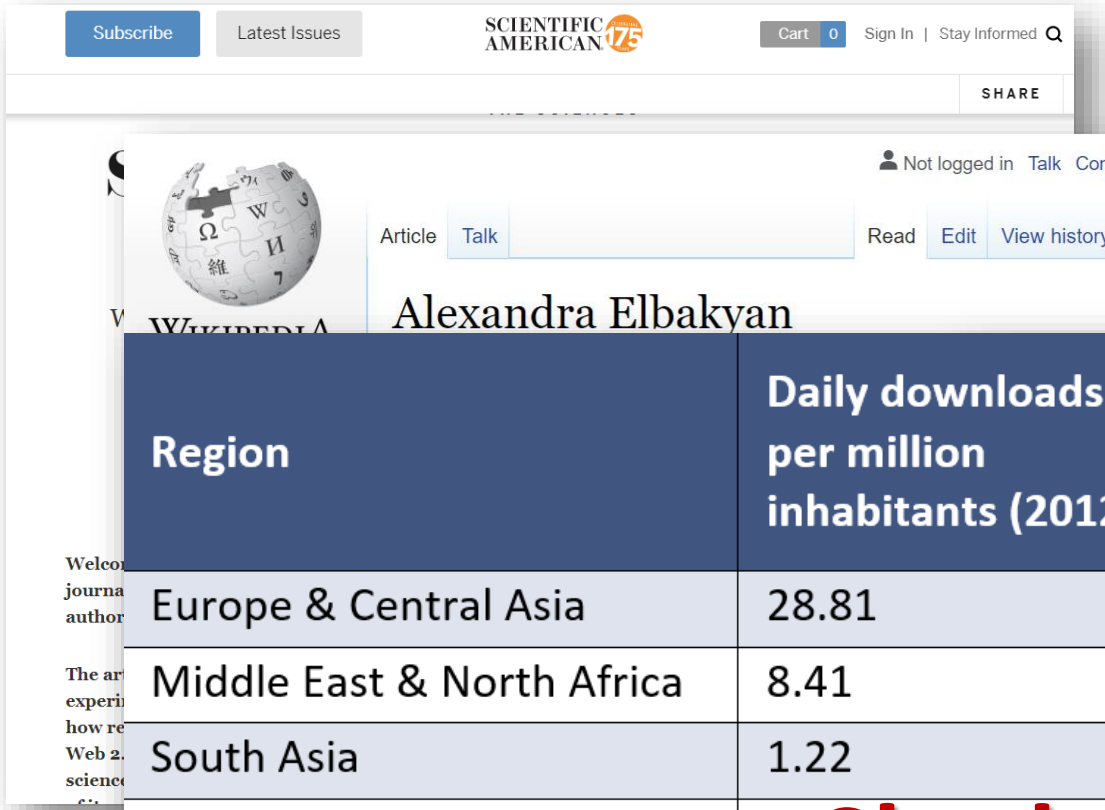
# sciencedirect.com



# link.springer.com



# Science 2.0 and Shadow Library

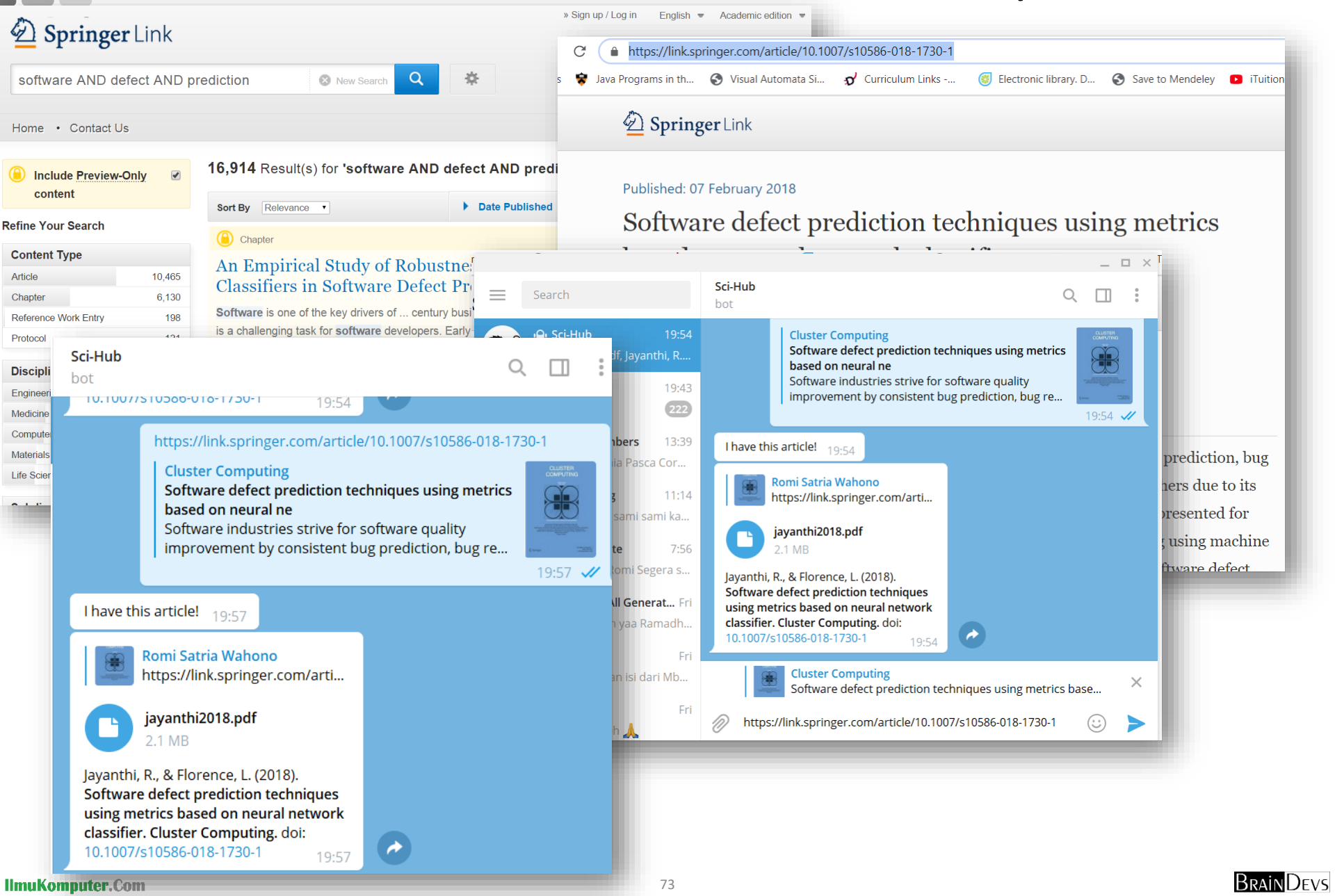



| Region | Daily downloads per million inhabitants (2012) | Daily downloads per million inhabitants (2015) | Growth (%) |
|---|---|---|---|
| Europe & Central Asia | 28.81 | 42.87 | 149% |
| Middle East & North Africa | 8.41 | 24.31 | 289% |
| South Asia | 1.22 | 6.38 | 523% |
| Latin America & Caribbean | 0.06 | 11.68 | 19467% |
| Sub-Saharan Africa | 0.28 | 3.98 | 1421% |
| East Asia & Pacific | 2.13 | [illegible] | 347% |
| North America | 10.47 | 42.76 | 408% |

**Shadow Library Statistics**

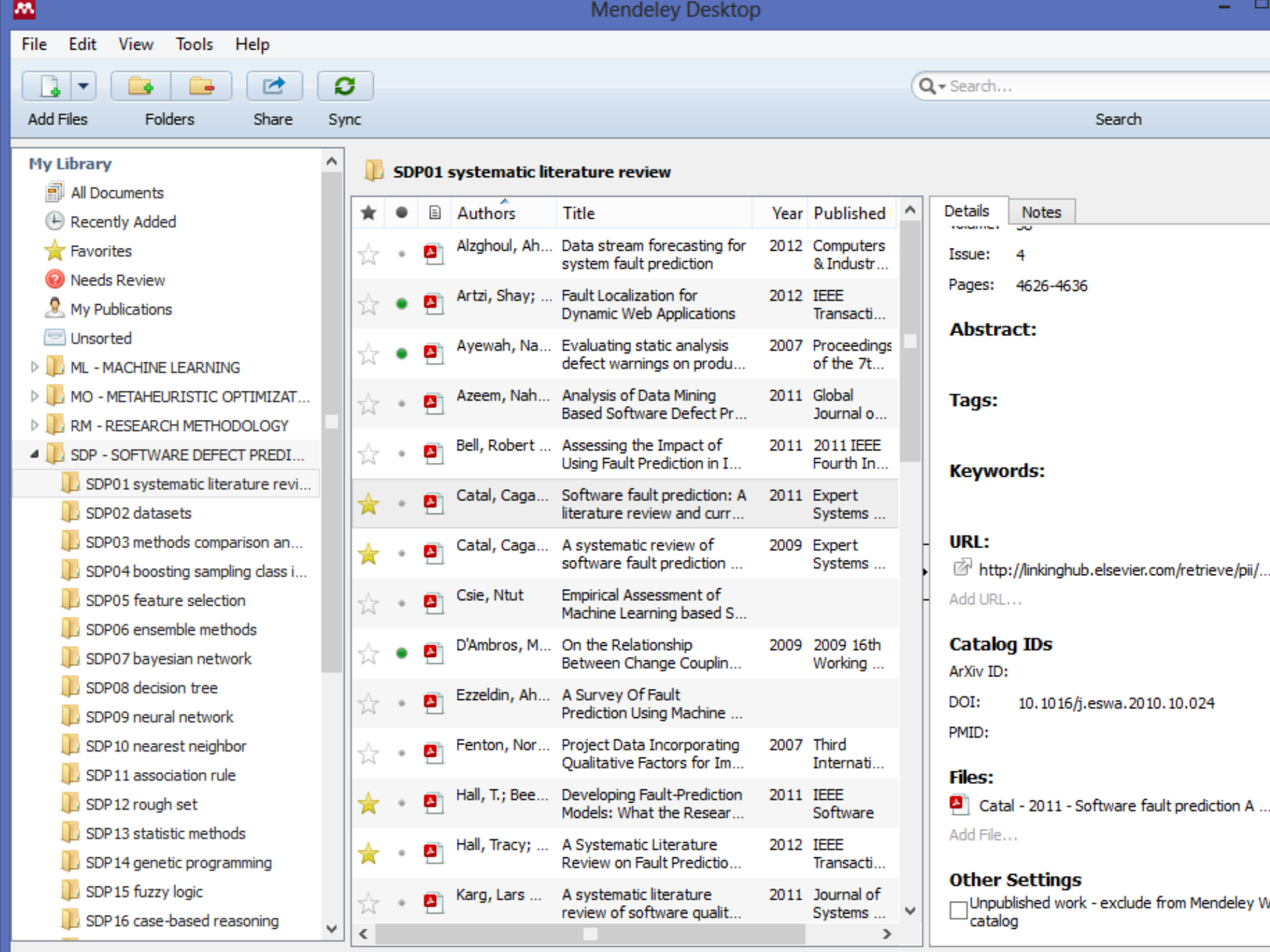
books. Using a dataset provided to us by one of the administrators of a prominent shadow library, we mapped both the supply of and the demand for academic monographs, textbooks and other learning materials via piratical shadow libraries. Our primary findings suggest that scholarly book piracy is a ubiquitous global phenomenon, with no apparent end in sight. If that is indeed true, what might the consequences be for the status quo in scholarly publishing?

# Science 2.0 and Shadow Library



# Metode Pengelolan Paper Yang Direview

- Buat account di mendeley.com dan download aplikasinya
- Buat folder dan sub folder pada aplikasi mendeley, dan letakkan paper (pdf) yang sudah didownload ke folder yang diinginkan di mendeley
- Bibliografi dari paper akan ditampilkan secara otomatis setelah pdf diletakkan
- Apabila mendeley gagal membaca bibliografi dari paper, cari DOI dari paper, isikan di catalog ID, pilih lookup, maka bibliografi akan muncul secara otomatis
- Install MS Word Plugin untuk membantu dalam menuliskan sitasi dan daftar referensi


Mendeley Desktop
File Edit View Tools Help
Add Files
Folders
Share
Sync
Search...
Search
My Library
All Documents
Recently Added
Favorites
Needs Review
My Publications
Unsorted
ML - MACHINE LEARNING
MO - METAHEURISTIC OPTIMIZAT...
RM - RESEARCH METHODOLOGY
SDP - SOFTWARE DEFECT PREDI...
SDP01 systematic literature revi...
SDP02 datasets
SDP03 methods comparison an...
SDP04 boosting sampling class i...
SDP05 feature selection
SDP06 ensemble methods
SDP07 bayesian network
SDP08 decision tree
SDP09 neural network
SDP10 nearest neighbor
SDP11 association rule
SDP12 rough set
SDP13 statistic methods
SDP14 genetic programming
SDP15 fuzzy logic
SDP16 case-based reasoning
SDP01 systematic literature review
Authors
Title
Year
Published
Alzghoul, Ah... Data stream forecasting for system fault prediction 2012 Computers & Industr...
Artzi, Shay; ... Fault Localization for Dynamic Web Applications 2012 IEEE Transacti...
Ayewah, Na... Evaluating static analysis defect warnings on produ... 2007 Proceedings of the 7t...
Azeem, Nah... Analysis of Data Mining Based Software Defect Pr... 2011 Global Journal o...
Bell, Robert ... Assessing the Impact of Using Fault Prediction in I... 2011 2011 IEEE Fourth In...
Catal, Caga... Software fault prediction: A literature review and curr... 2011 Expert Systems ...
Catal, Caga... A systematic review of software fault prediction ... 2009 Expert Systems ...
Csie, Ntut Empirical Assessment of Machine Learning based S...
D'Ambros, M... On the Relationship Between Change Couplin... 2009 2009 16th Working ...
Ezzeldin, Ah... A Survey Of Fault Prediction Using Machine ...
Fenton, Nor... Project Data Incorporating Qualitative Factors for Im... 2007 Third Internati...
Hall, T.; Bee... Developing Fault-Prediction Models: What the Resear... 2011 IEEE Software
Hall, Tracy; ... A Systematic Literature Review on Fault Predictio... 2012 IEEE Transacti...
Karg, Lars ... A systematic literature review of software qualit... 2011 Journal of Systems ...
Details
Notes
Issue: 4
Pages: 4626-4636
Abstract:
Tags:
Keywords:
URL:
http://linkinghub.elsevier.com/retrieve/pii/...
Add URL...
Catalog IDs
ArXiv ID:
DOI: 10.1016/j.eswa.2010.10.024
PMID:
Files:
Catal - 2011 - Software fault prediction A ...
Add File...
Other Settings
Unpublished work - exclude from Mendeley W
catalog

Expert Systems with Applications 38 (2011) 4626–4636

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Review

# Software fault prediction: A literature review and current trends

Cagatay Catal

*The Scientific and Technological Research Council of Turkey (TUBITAK), Marmara Research Center, Information Technologies Institute, Kocaeli, Turkey*

ARTICLE INFO

*Keywords:*
Machine learning
Automated fault prediction models
Expert systems
Software quality engineering
Software engineering
Statistical methods

ABSTRACT

Software engineering discipline contains several prediction approaches such as test effort prediction, correction cost prediction, fault prediction, reusability prediction, security prediction, effort prediction, and quality prediction. However, most of these prediction approaches are still in preliminary phase and more research should be conducted to reach robust models. Software fault prediction is the most popular research area in these prediction approaches and recently several research centers started new projects on this area. In this study, we investigated 90 software fault prediction papers published between year 1990 and year 2009 and then we categorized these papers according to the publication year. This paper surveys the software engineering literature on software fault prediction and both machine learning based and statistical based approaches are included in this survey. Papers explained in this article reflect the outline of what was published so far, but naturally this is not a complete review of all the papers published so far. This paper will help researchers to investigate the previous studies from metrics, methods, datasets, performance evaluation metrics, and experimental results perspectives in an easy and effective manner. Furthermore, current trends are introduced and discussed.



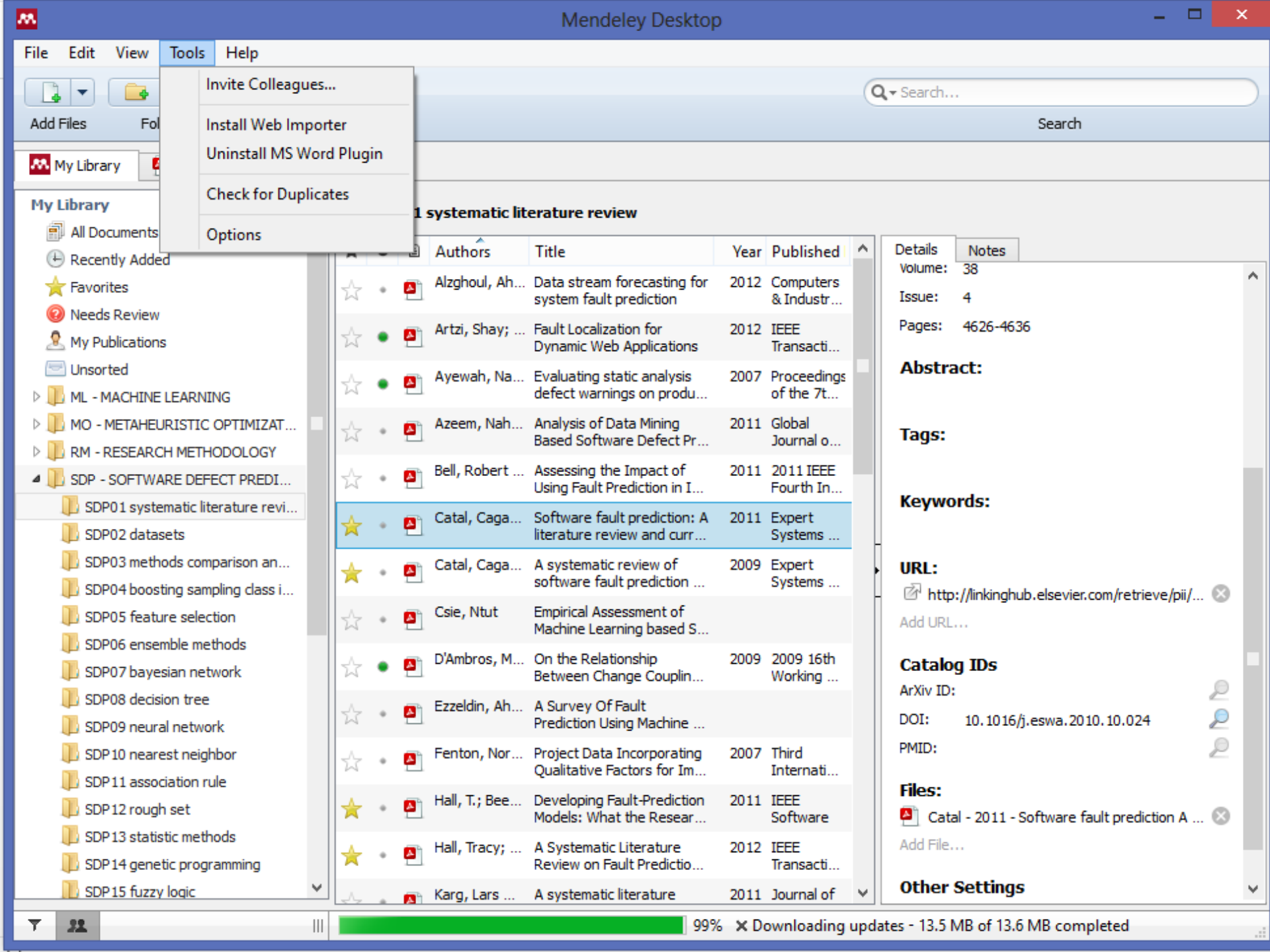
## 1. Introduction

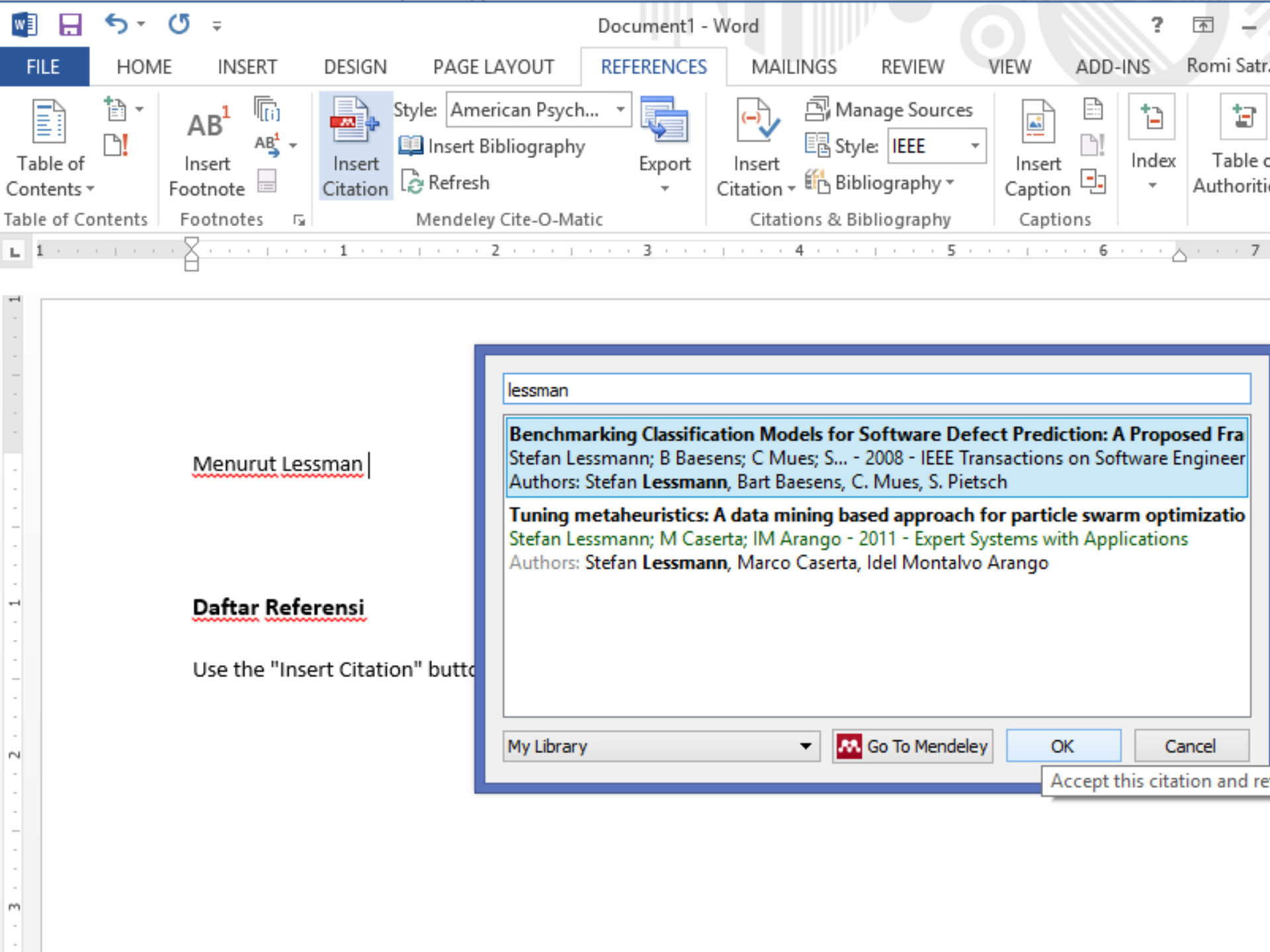
Software fault prediction approaches use previous software metrics and fault data to predict fault-prone modules for the next release of software. If an error is reported during system tests or from field tests, that module's fault data is marked as 1, otherwise 0. For prediction modeling, software metrics are used as *indepen-*

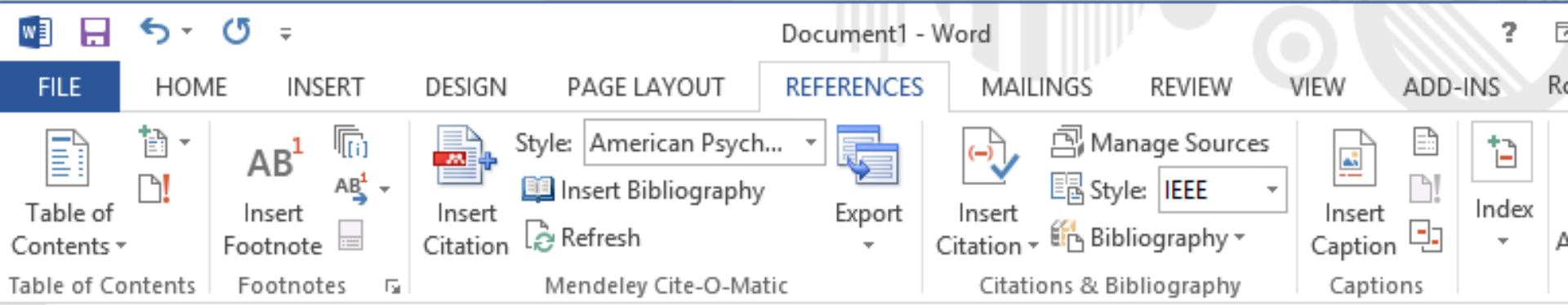
software fault prediction approaches are much more cost-effective to detect software faults compared to software reviews.

Benefits of software fault prediction are listed as follows (Catal & Diri, 2009a):

- Reaching a highly dependable system,
- Improving test process by focusing on fault-prone modules,


Mendeley Desktop
File
Edit
View
Tools
Help
Invite Colleagues...
Install Web Importer
Uninstall MS Word Plugin
Check for Duplicates
Options
Add Files
Search...
Search
My Library
My Library
All Documents
Recently Added
Favorites
Needs Review
My Publications
Unsorted
ML - MACHINE LEARNING
MO - METAHEURISTIC OPTIMIZAT...
RM - RESEARCH METHODOLOGY
SDP - SOFTWARE DEFECT PREDI...
SDP01 systematic literature revi...
SDP02 datasets
SDP03 methods comparison an...
SDP04 boosting sampling class i...
SDP05 feature selection
SDP06 ensemble methods
SDP07 bayesian network
SDP08 decision tree
SDP09 neural network
SDP10 nearest neighbor
SDP11 association rule
SDP12 rough set
SDP13 statistic methods
SDP14 genetic programming
SDP15 fuzzy logic
1 systematic literature review
Authors
Title
Year
Published
Alzghoul, Ah...
Data stream forecasting for system fault prediction
2012
Computers & Industr...
Artzi, Shay; ...
Fault Localization for Dynamic Web Applications
2012
IEEE Transacti...
Ayewah, Na...
Evaluating static analysis defect warnings on produ...
2007
Proceedings of the 7t...
Azeem, Nah...
Analysis of Data Mining Based Software Defect Pr...
2011
Global Journal o...
Bell, Robert ...
Assessing the Impact of Using Fault Prediction in I...
2011
2011 IEEE Fourth In...
Catal, Caga...
Software fault prediction: A literature review and curr...
2011
Expert Systems ...
Catal, Caga...
A systematic review of software fault prediction ...
2009
Expert Systems ...
Csie, Ntut
Empirical Assessment of Machine Learning based S...
D'Ambros, M...
On the Relationship Between Change Couplin...
2009
2009 16th Working ...
Ezzeldin, Ah...
A Survey Of Fault Prediction Using Machine ...
Fenton, Nor...
Project Data Incorporating Qualitative Factors for Im...
2007
Third Internati...
Hall, T.; Bee...
Developing Fault-Prediction Models: What the Resear...
2011
IEEE Software
Hall, Tracy; ...
A Systematic Literature Review on Fault Predictio...
2012
IEEE Transacti...
Karg, Lars ...
A systematic literature
2011
Journal of
Details
Notes
Volume: 38
Issue: 4
Pages: 4626-4636
Abstract:
Tags:
Keywords:
URL:
http://linkinghub.elsevier.com/retrieve/pii/...
Add URL...
Catalog IDs
ArXiv ID:
DOI: 10.1016/j.eswa.2010.10.024
PMID:
Files:
Catal - 2011 - Software fault prediction A ...
Add File...
Other Settings
99%
Downloading updates - 13.5 MB of 13.6 MB completed


Document1 - Word
FILE
HOME
INSERT
DESIGN
PAGE LAYOUT
REFERENCES
MAILINGS
REVIEW
VIEW
ADD-INS
Romi Satr
Table of Contents
Insert Footnote
Insert Citation
Style: American Psych...
Insert Bibliography
Refresh
Export
Insert Citation
Manage Sources
Style: IEEE
Bibliography
Insert Caption
Index
Table of Contents
Footnotes
Mendeley Cite-O-Matic
Citations & Bibliography
Captions
Menurut Lessman
Daftar Referensi
Use the "Insert Citation" butt
lessman
Benchmarking Classification Models for Software Defect Prediction: A Proposed Fra
Stefan Lessmann; B Baesens; C Mues; S... - 2008 - IEEE Transactions on Software Engineer
Authors: Stefan Lessmann, Bart Baesens, C. Mues, S. Pietsch
Tuning metaheuristics: A data mining based approach for particle swarm optimizatio
Stefan Lessmann; M Caserta; IM Arango - 2011 - Expert Systems with Applications
Authors: Stefan Lessmann, Marco Caserta, Idel Montalvo Arango
My Library
Go To Mendeley
OK
Cancel
Accept this citation and re




Menurut Lessman (Lessmann, Baesens, Mues, & Pietsch, 2008), prediksi cacat software saat ini

**Daftar Referensi**

Lessmann, S., Baesens, B., Mues, C., & Pietsch, S. (2008). Benchmarking Classification Models for Software Defect Prediction: A Proposed Framework and Novel Findings. *IEEE Transactions on Software Engineering*, *34*(4), 485–496. doi:10.1109/TSE.2008.35

# Jenis Paper Ilmiah

1. Technical Paper
    1. Paper yang isinya adalah hasil penelitian dan eksperimen yang dilakukan seorang peneliti
    2. Penilaian kualitas technical paper dari kontribusi ke pengetahuan

2. Survey Paper
    1. Paper yang isinya adalah review dan survey tentang topik/tema suatu penelitian, biasanya jumlah penelitian yang direview mencapai ratusan atau ribuan paper
    2. Rujukan dan panduan penting bagi peneliti yang baru memulai penelitian untuk memahami suatu topic/tema penelitian secara komprehensif

# Kiat Mereview Technical Paper

1. Pahami Masalah Penelitian
   - Apakah penelitian hanya menyelesaikan masalah yang dibuat-buat?
   - Apakah masalah penelitian dilandasi dan divalidasi?
2. Pahami Kontribusi
   - Apakah peneliti hanya mengulang hal yang sudah ada?
   - Apakah peneliti menyadari literatur lain yang berhubungan dengan penelitiannya?
   - Apa yang baru dan orisinil di paper itu (metodologi, algoritma, evaluasi, validasi, tool, dsb.)?
3. Pahami Validitas Kontribusi
   - Apakah teori atau model yang diusulkan sudah terbukti benar? Tidak adakah kesalahan pada pembuktian?
   - Adakah faktor-faktor aneh pada proses eksperimen penelitian?
   - Apakah benchmark yang dilakukan realistis atau hanya buatan? Ataukah membandingkan apel dan jeruk?
   - Apakah generalisasi cukup valid?

# Kiat Mereview Paper Survey

- Secara umum, paper survey harus dibaca seluruh bagian untuk mendapatkan pemahaman yang komprehensif

- Fokus ke bagian paper (tergantung jenis paper survey): Traditional Review or Systematic Literature Review (SLR)
  1. Traditional Review:
     - Lihat bagian yang menjelaskan tentang state-of-the-art methods
  2. Systematic Literature Review (SLR)
     - Pahami Research Question (RQ) yang biasanya tertulis secara eksplisit di paper, dimana Jawaban RQ ada di bagian "result and analysis"
     - Perhatikan RQ tentang:
       - "best model/method/algorithm" karena di situ akan dibahas tentang state-of-the-art method
       - "research challenge/problems", karena di situ kita bisa menemukan masalah penelitian terkini (state-of-the-art problem)

# Jenis dan Metode Penulisan Survey Paper

1. Traditional Review

2. Systematic Review
    1. Systematic Mapping Study (Scoping Study)
    2. Systematic Literature Review (SLR)
    3. Tertiary Study

# 1. Traditional Review

- Provides an overview of the research findings on particular topics
- Advantages: produce insightful, valid syntheses of the research literature if conducted by the expert
- Disadvantages: vulnerable to unintentional and intentional bias in the selection, interpretation and organization of content
- Examples:
  - Liao et al., Intrusion Detection System: A Comprehensive Review, Journal of Network and Computer Applications, 36(2013)
  - Galar et al., A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), Vol. 42, No. 4, July 2012
  - Cagatay Catal, Software fault prediction: A literature review and current trends, Expert Systems with Applications 38 (2011)

# 2. Systematic Mapping Study

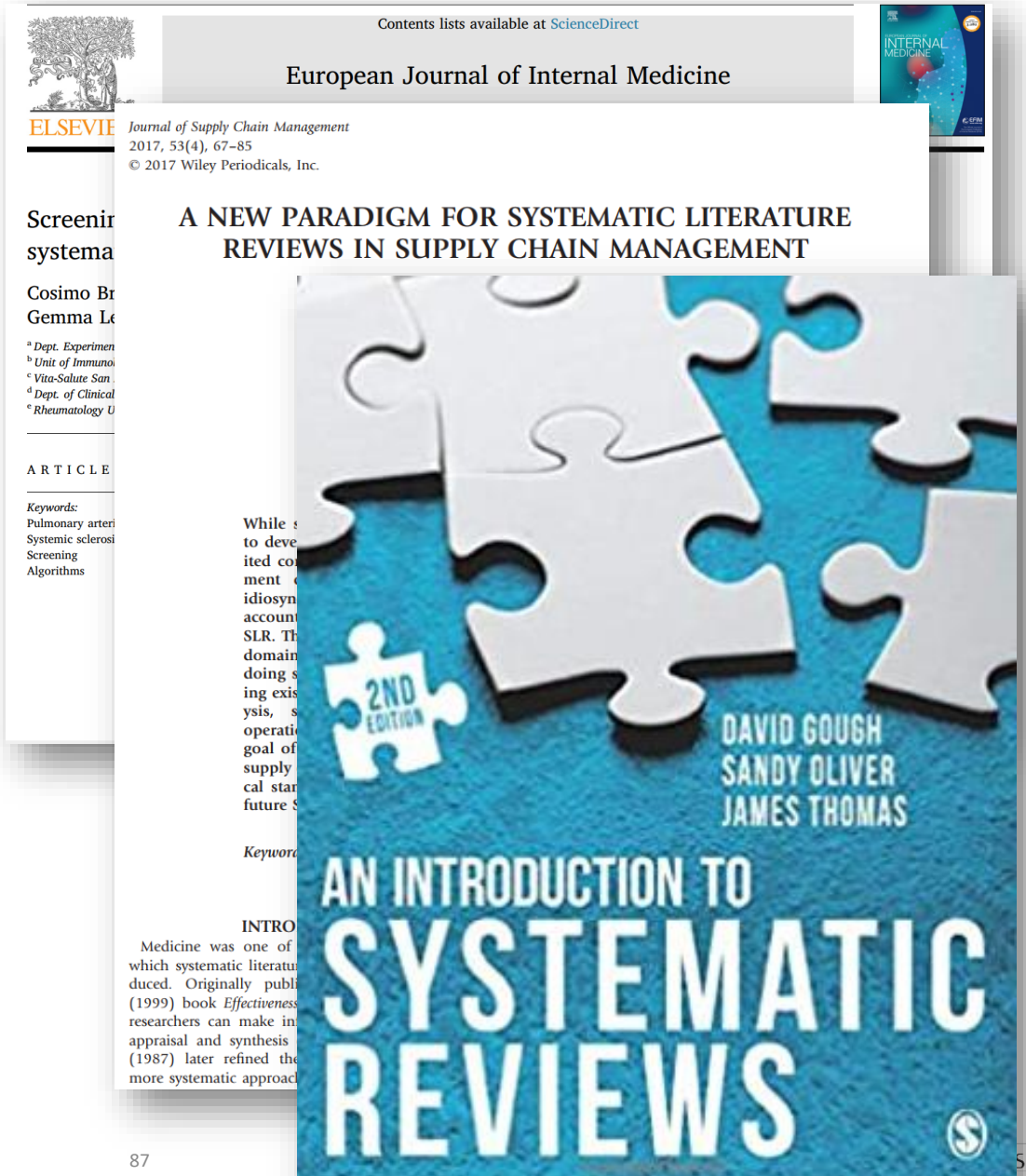
- Suitable for a very broad topic
- Identify clusters of evidence (making classification)
- Direct the focus of future SLRs
- To identify areas for future primary studies
- Examples:
  - Neto et al., A systematic mapping study of software product lines testing, Information and Software Technology Vol. 53, Issue 5, May 2011
  - Elberzhager et al., Reducing test effort: A systematic mapping study on existing approaches, Information and Software Technology 54 (2012)

# 3. Systematic Literature Review (SLR)

- Systematic reviews are a type of literature review that uses systematic methods to collect secondary data, critically appraise research studies, and synthesize findings qualitatively or quantitatively *(Amstrong et al., 2011)*
- A process of identifying, assessing, and interpreting all available research evidence, to provide answers for a particular research question (RQ)
- They are designed to provide a complete, exhaustive summary of current evidence, that is methodical, comprehensive, transparent, and replicable
- SLRs are well established in other disciplines, particularly medicine, biomedic, healthcare
- SLR application in the various fields:
  - Medicine *(Archie Cochrane, 1974)*
  - Computing Field *(Kitchenham & Charter, 2007)*
  - Social Science *(Gough, 2016)*
  - Business Management *(Durach et al., 2017)*

# Contoh Systematic Literature Review (SLR)

- Romi Satria Wahono, A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks, Journal of Software Engineering, Vol. 1, No. 1, April 2015
- Christian F. Durach  Joakim Kembro  Andreas Wieland, A New Paradigm for Systematic Literature Reviews in Supply Chain Management, Journal of Supply Chain Management, Vol. 53(4), pp 67–85, 2017
- Matthias Galster, Danny Weyns, Dan Tofan, Bartosz Michalik, and Paris Avgeriou, Variability in Software Systems: A Systematic Literature Review, IEEE Transactions on Software Engineering, Vol 40, No 3, 2014

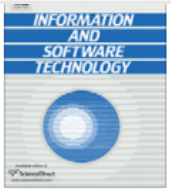


# Contoh Systematic Literature Review (SLR)

Contents lists available at SciVerse ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof




Systematic literature review of machine learning based software development effort estimation

Jianfeng Wen [a,*], Shix

[a] Department of Computer Science,
[b] Department of Computer Science,
[c] Institute of Business Intelligence a
[d] Engineering Research Center of Co

ARTICLE INFO

*Article history:*
Received 28 October 2010
Received in revised form 8 Augus
Accepted 8 September 2011
Available online 16 September 20

*Keywords:*
Software effort estimation
Machine learning
Systematic literature review

IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 38, NO. 6, NOVEMBER/DECEMBER 2

# A Systematic Literature Review on Fault Prediction Performance in Software Engineering

Tracy Hall, Sarah Beecham, David Bowes, David Gray, and Steve Counsell

**Abstract**—*Background*: The accurate prediction of where faults are likely to occur in code can help direct test effort, reduce costs, and improve the quality of software. *Objective*: We investigate how the context of models, the independent variables used, and the modeling techniques applied influence the performance of fault prediction models. *Method:* We used a systematic literature review to identify 208 fault prediction studies published from January 2000 to December 2010. We synthesize the quantitative and qualitative results of 36 studies which report sufficient contextual and methodological information according to the criteria we develop and apply. *Results:* The models that perform well tend to be based on simple modeling techniques such as Naive Bayes or Logistic Regression. Combinations of independent variables have been used by models that perform well. Feature selection has been applied to these combinations when models are performing particularly well. *Conclusion:* The methodology used to build models seems to be influential to predictive performance. Although there are a set of fault prediction studies in which confidence is possible, more studies are needed that use a reliable methodology and which report their context, methodology, and performance comprehensively.

# 4. Tertiary study

- Is a SLR of SLRs
- To answer a more wider question
- Uses the same method as in SLR
- Potentially less resource intensive

- Examples:
  - Kitchenham et al., Systematic literature reviews in software engineering – A tertiary study, *Information and Software Technology 52 (2010)*
  - Cruzes et al., Research synthesis in software engineering: A tertiary study, *Information and Software Technology 53 (2011)*

# 3. Systematic Literature Review (SLR)

3.1 Tahapan SLR
3.2 Contoh SLR

# 3.1 Tahapan SLR

# Tahapan SLR

1. Formulate the review's research question
2. Develop the review's protocol

PLANNING

1. Identify the relevant literature
2. Perform selection of primary studies
3. Perform data extraction
4. Assess studies' quality
5. Conduct synthesis of evidence

CONDUCTING

1. Write up the SLR report/paper
2. Choose the Right Journal

REPORTING

# 1. Tahapan Planning

1. Formulate the Review's Research Question
2. Develop the Review's Protocol

# 1. Formulate the Review's Research Question

- Features of good question:
  - The RQ is meaningful and important to practitioners and researchers.
  - The RQ will lead to changes in current software engineering practice or to increase confidence in the value of current practice
  - The RQ will identify discrepancies between commonly held beliefs and the reality
- RQ can be derived primarily based on researcher's interest
  - An SLR for PhD thesis should identify existing basis for the research work and where it fits in the current body of knowledge

# The Research Question (RQ)

- Is the most important part in any SLR
- Is not necessarily the same as questions addressed in your research
- Is used to guide the search process
- Is used to guide the extraction process
- Data analysis (synthesis of evidence) is expected to answer your SLR's RQ

# RQ and PICOC

The formulation of RQs about effectiveness of a treatment should focus on 5 elements known as PICOC:

1. **Population (P)** - the target group for the investigation (e.g. people, software etc.)
2. **Intervention (I)** - specifies the investigation aspects or issues of interest to the researchers
3. **Comparison (C)**– aspect of the investigation with which the intervention is being compared to
4. **Outcomes (O)**– the effect of the intervention
5. **Context (C)**– the setting or environment of the investigation

*(Petticrew et al., Systematic Reviews in the Social Sciences: A Practical Guide, Blackwell Publishing, 2006)*

# Example of PICOC (Kitchenham et al., 2007)

Kitchenham et al., A Systematic Review of Cross- vs. Within-Company Cost Estimation Studies, *IEEE Transactions on Software Engineering*, 33 (5), 2007

| | |
|---|---|
| Population: | Software or web project |
| Intervention: | Cross-company project effort estimation model |
| Comparison: | Single-company project effort estimation model |
| Outcomes: | Prediction or estimate accuracy |
| Context: | None |

# Example of PICOC (Wahono, 2015)

Romi Satria Wahono, A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks, *Journal of Software Engineering, Vol. 1, No. 1*, pp. 1-16, April 2015

| | |
|---|---|
| **Population** | Software, software application, software system, information system |
| **Intervention** | Software defect prediction, fault prediction, error-prone, detection, classification, estimation, models, methods, techniques, datasets |
| **Comparison** | n/a |
| **Outcomes** | Prediction accuracy of software defect, successful defect prediction methods |
| **Context** | Studies in industry and academia, small and large data sets |

# Example of RQs (Kitchenham, 2007)

Kitchenham et al., A Systematic Review of Cross- vs. Within-Company Cost Estimation Studies, *IEEE Transactions on Software Engineering*, 33 (5), 2007

- RQ1: What evidence is there that cross-company estimation models are not significantly different from within-company estimation models for predicting effort for software/Web projects?
- RQ2: What characteristics of the study data sets and the data analysis methods used in the study affect the outcome of within- and cross-company effort estimation accuracy studies?
- RQ3: Which experimental procedure is most appropriate for studies comparing within- and cross-company estimation models?

# Example of RQs (Radjenovic et al., 2013)

Radjenovic et al., Software fault prediction metrics: A systematic literature review, *Information and Software Technology,* Vol. 8, No. 55, pp. 1397-1418, 2013

- RQ1: Which software metrics for fault prediction exist in literature?
- RQ2: What data sets are used for evaluating metrics?

# Example of RQs (Fu Jia et al., 2020)

Fu Jia et al., Soybean Supply Chain Management and Sustainability: A Systematic Literature Review, *Journal of Cleaner Production*, 2020

- RQ1: What are the drivers and barriers to sustainable soy production and their relationships?
- RQ2: What are the value chain governance mechanisms available for the soybean chain?
- RQ3: What are the consequences of the implementation of these mechanisms?

# Example of RQ (Wahono, 2015)

Romi Satria Wahono, A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks, *Journal of Software Engineering, Vol. 1, No. 1*, pp. 1-16, April 2015

| ID | Research Question |
|---|---|
| RQ1 | Which journal is the most significant software defect prediction journal? |
| RQ2 | Who are the most active and influential researchers in the software defect prediction field? |
| RQ3 | What kind of research topics are selected by researchers in the software defect prediction field? |
| RQ4 | What kind of datasets are the most used for software defect prediction? |
| RQ5 | What kind of methods are used for software defect prediction? |
| RQ6 | What kind of methods are used most often for software defect prediction? |
| RQ7 | Which method performs best when used for software defect prediction? |
| RQ8 | What kind of method improvements are proposed for software defect prediction? |
| RQ9 | What kind of frameworks are proposed for software defect prediction? |

# 2. Develop the Review's Protocol

- A plan that specifies the basic review procedures (method)
- Components of a protocol:
  1. Background
  2. Research Questions
  3. Search terms
  4. Selection criteria
  5. Quality checklist and procedures
  6. Data extraction strategy
  7. Data synthesis strategy

# 2. Tahapan Conducting

1. Identify the Relevant Literature
2. Perform Selection of Primary Studies
3. Perform Data Extraction
4. Assess Studies' Quality
5. Conduct Synthesis of Evidence

# 1. Identifying Relevant Literature

- A comprehensive and exhaustive searching of studies to be included in the review
- Define a search strategy
- Search strategies are usually iterative and benefit from:
  - Preliminary searches (to identify existing review and volume of studies)
  - Trial searches (combination of terms from RQ)
  - Check the search results against list of known studies
  - Consult the experts in the field

# Approach to Construct Search String

- Derive major terms used in the review questions based on the PICOC
- List the keywords mentioned in the article
- Search for synonyms and alternative words
- Use the boolean OR to incorporate alternative synonyms
- Use the boolean AND to link major terms

# Example of Search String (Kitchenham et al., 2007)

- Kitchenham et al. (2007) used their structured questions to construct search strings for use with electronic databases:
  - *Population*: software OR application OR product OR Web OR WWW OR Internet OR World-Wide Web OR project OR development
  - *Intervention*: cross company OR cross organisation OR cross organization OR multiple-organizational OR multiple-organisational model OR modeling OR modelling effort OR cost OR resource estimation OR prediction OR assessment
  - *Contrast*: within-organisation OR within-organization OR within-organizational OR within-organisational OR single company OR single organisation
  - *Outcome*: Accuracy OR Mean Magnitude Relative Error
- The search strings were constructed by linking the four OR lists using the Boolean AND

# Example of Search String (Wahono, 2015)

Romi Satria Wahono, A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks, *Journal of Software Engineering, Vol. 1, No. 1*, pp. 1-16, April 2015

Search String:

*(software OR applicati* OR systems ) AND*

*(fault* OR defect* OR quality OR error-prone) AND*

*(predict* OR prone* OR probability OR assess* OR detect* OR estimat* OR classificat*)*

# Example of Search String (Salleh et al., 2011)

- The complete search term initially used:

  *(student* OR undergraduate*) AND (pair programming OR pair-programming) AND ((experiment* OR measurement OR evaluation OR assessment) AND (effective* OR efficient OR successful)*

- A very limited number of results retrieved when using the complete string
  - thus a much simpler string was derived
- Subject librarian suggested to revise the search string:

*"pair programming" OR "pair-programming"*

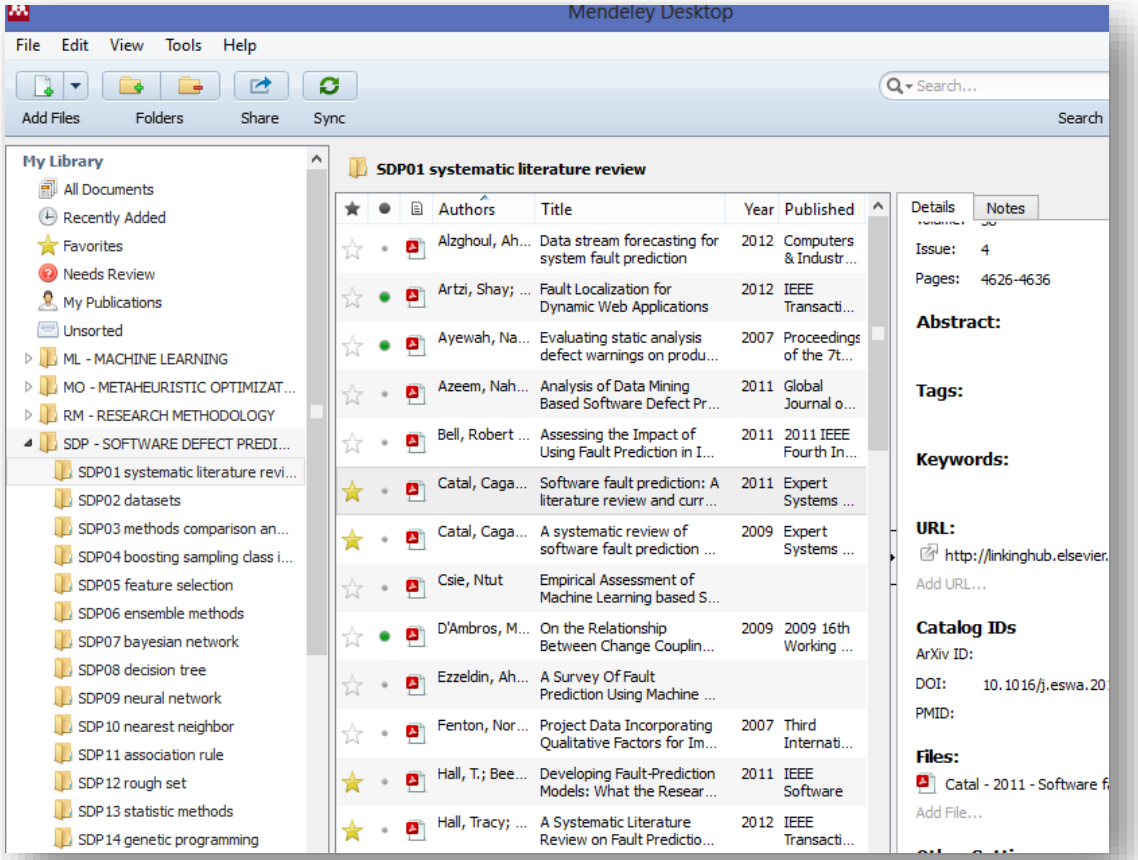
# Sources of Evidence (Kitchenham et al., 2007)

- The search strings were used on 6 digital libraries:
  - Science Direct, SpringerLink, SCOPUS, Web of Science, IEEExplore, ACM Digital library

- Search specific journals and conf. proceedings:
  - Empirical Software Engineering (J)
  - Information and Software Technology (J)
  - Software Process Improvement and Practice (J)
  - International Conference on Software Engineering (C)
  - Journal of Business Research (J)
  - Management Science (J)
  - International Business Review (J)

# Managing Bibliography

- Use relevant Bibliographic package to manage large number of references
- E.g. Mendeley, EndNote, Zotero, JabRef Reference Manager etc.



# Documenting the Search

- The process of conducting SLR must be transparent and replicable
- The review should be documented in sufficient detail
- The search should be documented and changes noted

| Data Source | Documentation |
|---|---|
| Digital Library | Name of Database, Search strategy, Date of search, years covered by search |
| Journal Hand Searches | Name of journal, Years searched |
| Conference proceedings | Title of proceedings/Name of conference, Journal name (if published as part of a journal) |

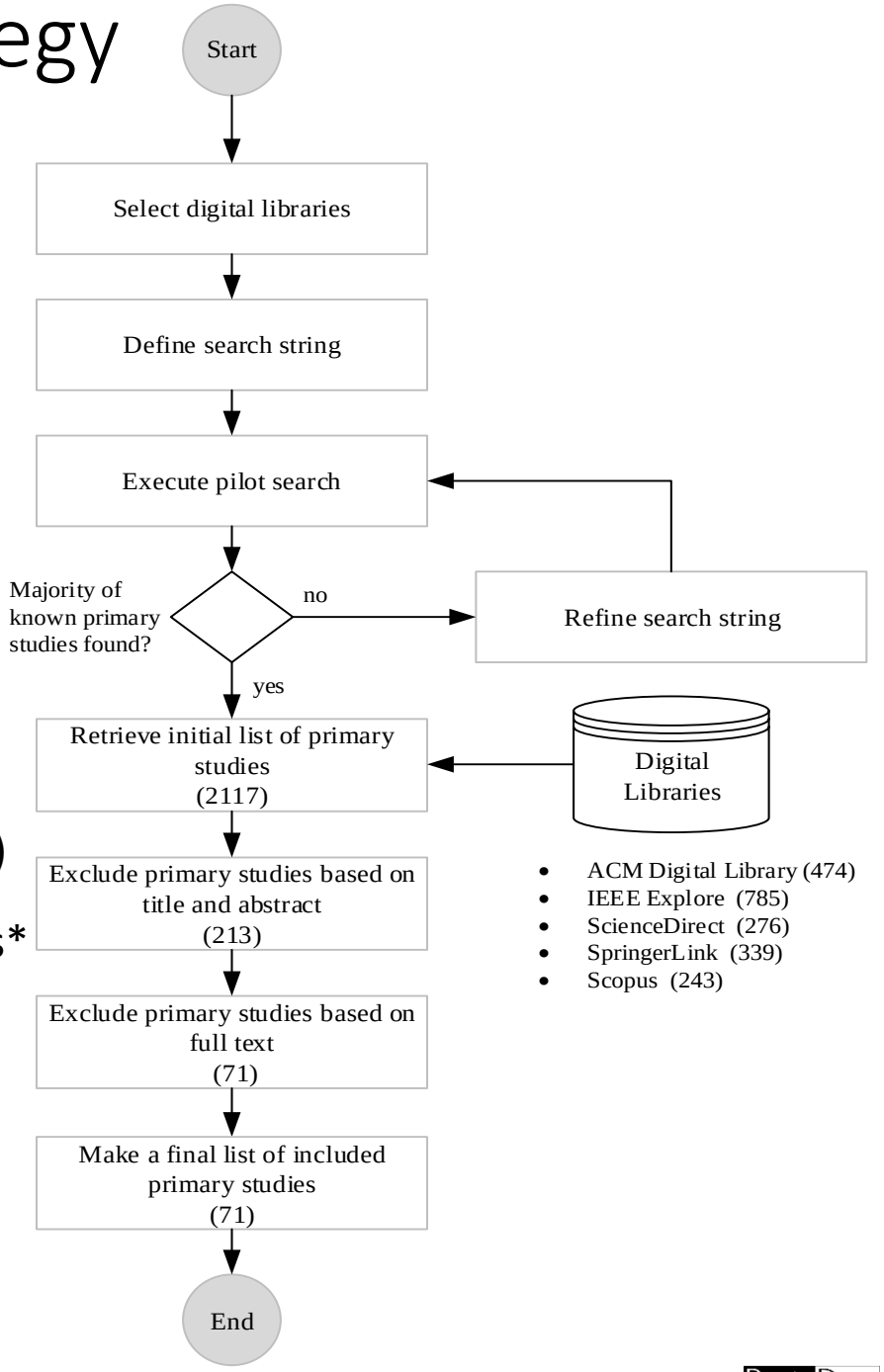
# 2. Selection of Studies

- Primary studies need to be assessed for their actual relevance
- Set the criteria for including or excluding studies (decided earlier during protocol development, can be refined later)
- Inclusion & exclusion criteria should be based on RQ
- Study selection is a multistage process

# Studies Selection Strategy (Wahono, 2015)

- Publication Year:
  - ✓2000-2013
- Publication Type:
  - ✓Journal
  - ✓Conference Proceedings
- Search String:

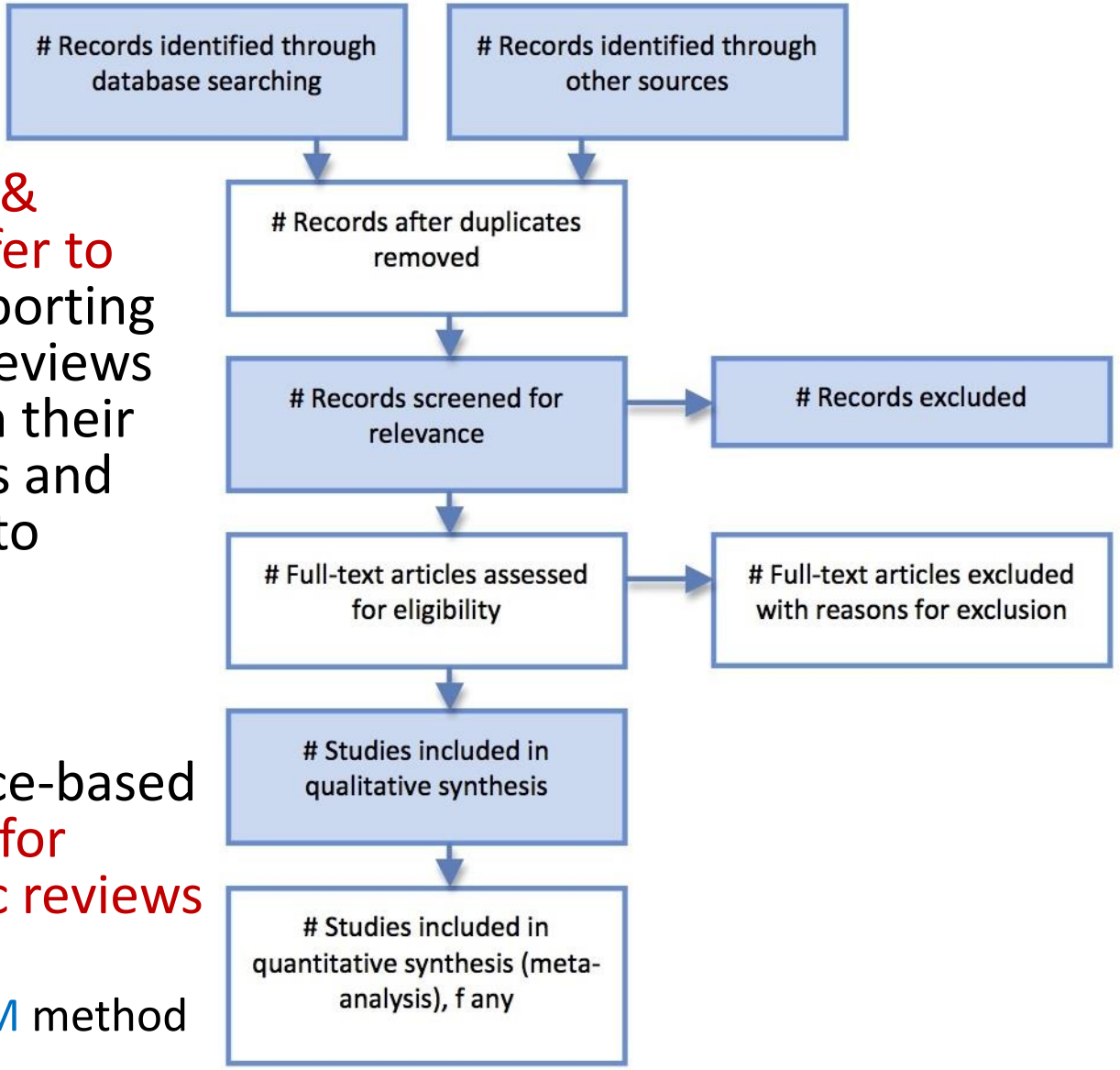
  software
  AND
  (fault* OR defect* OR quality OR error-prone)
  AND
  (predict* OR prone* OR probability OR assess* OR detect* OR estimat* OR classificat*)
- Selected Studies:
  - ✓71



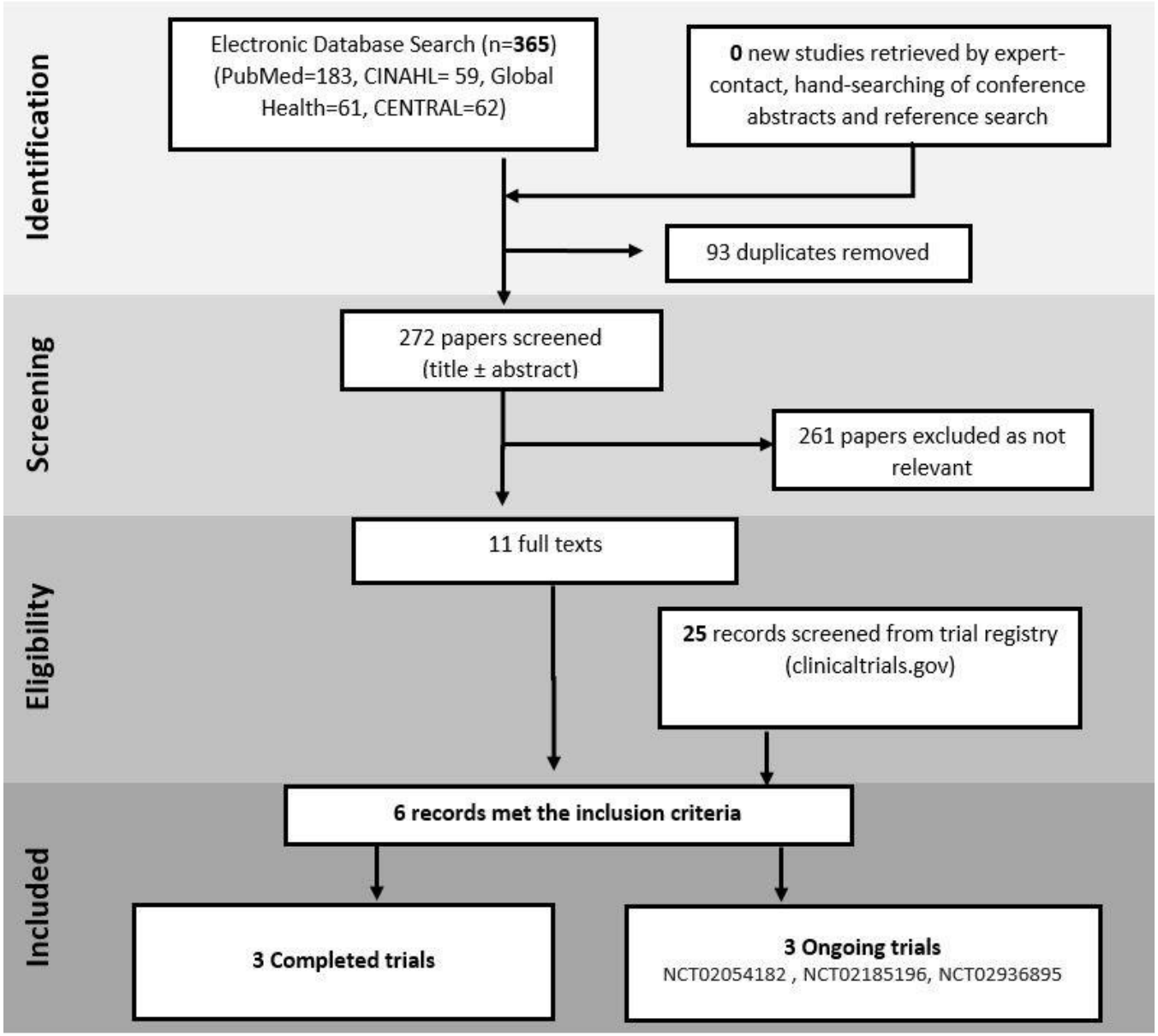
# Studies Selection Strategy (PRISMA)

- Many leading medical & healthcare journals refer to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses ) in their Instructions to Authors and some require authors to adhere to them *(Liberati et al., 2009)*

- (PRISMA) is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses
  - Replaced the QUOROM method



# Studies Selection Strategy (PRISMA)



# Selection of Studies (Kitchenham et al., 2007)

- Kitchenham et al. (2007) used the following inclusion criteria:
  - Any study that compared predictions of cross-company models with within-company models based on analysis of single company project data

- They used the following exclusion criteria:
  - Studies where projects were only collected from a small number of different sources (e.g. 2 or 3 companies)
  - Studies where models derived from a within-company data set were compared with predictions from a general cost estimation model

# Selection of Studies (Wahono, 2015)

| | |
|---|---|
| **Inclusion Criteria** | Studies in academic and industry using large and small scale data sets |
| | Studies discussing and comparing modeling performance in the area of software defect prediction |
| | For studies that have both the conference and journal versions, only the journal version will be included |
| | For duplicate publications of the same study, only the most complete and newest one will be included |
| **Exclusion Criteria** | Studies without a strong validation or including experimental results of software defect prediction |
| | Studies discussing defect prediction datasets, methods, frameworks in a context other than software defect prediction |
| | Studies not written in English |

# Selection of Studies (*Salleh et al., 2011)*

- **Inclusion criteria**:
  - to include any empirical studies of PP that involved higher education students as the population of interest

- **Exclusion criteria**:
  - Papers presenting unsubstantiated claims made by the author(s), for which no evidence was available.
  - Papers about Agile/XP describing development practices other than PP, such as test-first programming, refactoring etc.
  - Papers that only described tools (software or hardware) that could support the PP practice.
  - Papers not written in English.
  - Papers involving students but outside higher education

# 3. Assessing Studies' Quality

- To provide more detailed Inclusion/Exclusion criteria
- To check whether quality differences provide an explanation for differences in study results
- As a means of weighting the importance of individual studies when results are being synthesized
- To guide the interpretation of findings and determine the strength of inferences
- To guide recommendations for further research

# Assessing Studies' Quality

- Quality relates to the extent to which the study:
  - Minimizes bias, and
  - Maximizes internal and external validity (Khan et al. 2001)
- Quality Concepts Definition (Kitchenham & Charter, 2007)

| Terms | Synonyms | Definition |
|---|---|---|
| Bias | Systematic error | tendency to produce results that depart systematically from the 'true' results. Unbiased results are internally valid |
| Internal Validity | Validity | The extent to which the design and conduct of the study are likely to prevent systematic error. Internal validity is a prerequisite for external validity |
| External Validity | Generalizability, Applicability | The extent to which the effects observed in the study are applicable outside of the study |

# Assessing Studies' Quality

- Assessing quality of studies:
  - Methodology or design of the study
  - Analysis of studies' findings
- Quality checklist or instrument need to be designed to facilitate quality assessment
- Most quality checklists include questions aimed at assessing the extent to which articles have addressed bias and validity

# Study Quality Assessment (Salleh et al., 2011)

| Item | Answer |
|---|---|
| 1. Was the article referred? [30] | Yes/No |
| 2. Were the aim(s) of the study clearly stated? [16], [67] | Yes/No/Partially |
| 3. Were the study participants or observational units adequately described? For example, students' programming experience, year of study etc. [44], [68] | Yes/No/Partially |
| 4. Were the data collections carried out very well? For example, discussion of procedures used for collection, and how the study setting may have influenced the data collected [44], [48], [67], [68] | Yes/No/Partially |
| 5. Were potential confounders adequately controlled for in the analysis? 67] | Yes/No/Partially |
| 6. Were the approach to and formulation of the analysis well conveyed? For example, description of the form of the original data, rationale for choice of method/tool/package [48], [67], [68] | Yes/No/Partially |
| 7. Were the findings credible? For example, the study was methodologically explained so that we can trust the findings; findings/conclusions are resonant with other knowledge and experience [48], [44], [68] | Yes/No/Partially |

# Study Quality Assessment (Kitchenham et al., 2007)

Kitchenham et al. (2007) constructed a quality questionnaire based on 5 issues affecting the quality of the study:

1. Is the data analysis process appropriate?
2. Did studies carry out a sensitivity or residual analysis?
3. Were accuracy statistics based on the raw data scale?
4. How good was the study comparison method?
5. The size of the within-company data set
   (e.g < 10 projects considered poor quality)

# 4. Data Extraction

- Involve reading the full text article
- Data extracted from primary studies should be recorded using *data extraction form*
- The form should be designed and piloted when the protocol is defined
- Collect all the information that can be used to answer the RQ and the study's quality criteria
- Both quality checklist and review data can be included in the same form
- In case of duplicates publications (reporting the same data), refer the most complete one
- For validation, a set of papers should be reviewed by 2 or more researchers. Compare results and resolve any conflicts

# 5. Synthesis of Evidence

- Involves collating and summarizing the results of the included primary studies
- Key objectives of data synthesis (Cruzes & Dyba, 2011):
  - to analyze and evaluate multiple studies
  - to select appropriate methods for integrating or providing new interpretive explanations about them
- Synthesis can be:
  - Descriptive (narrative/non-quantitative)
  - Quantitative (e.g. meta-analysis)

(Cruzes et al., Research Synthesis in Software Engineering: A tertiary study, *Information and Software Technology*, 53(5), 2011)

# Descriptive Synthesis (Narrative)

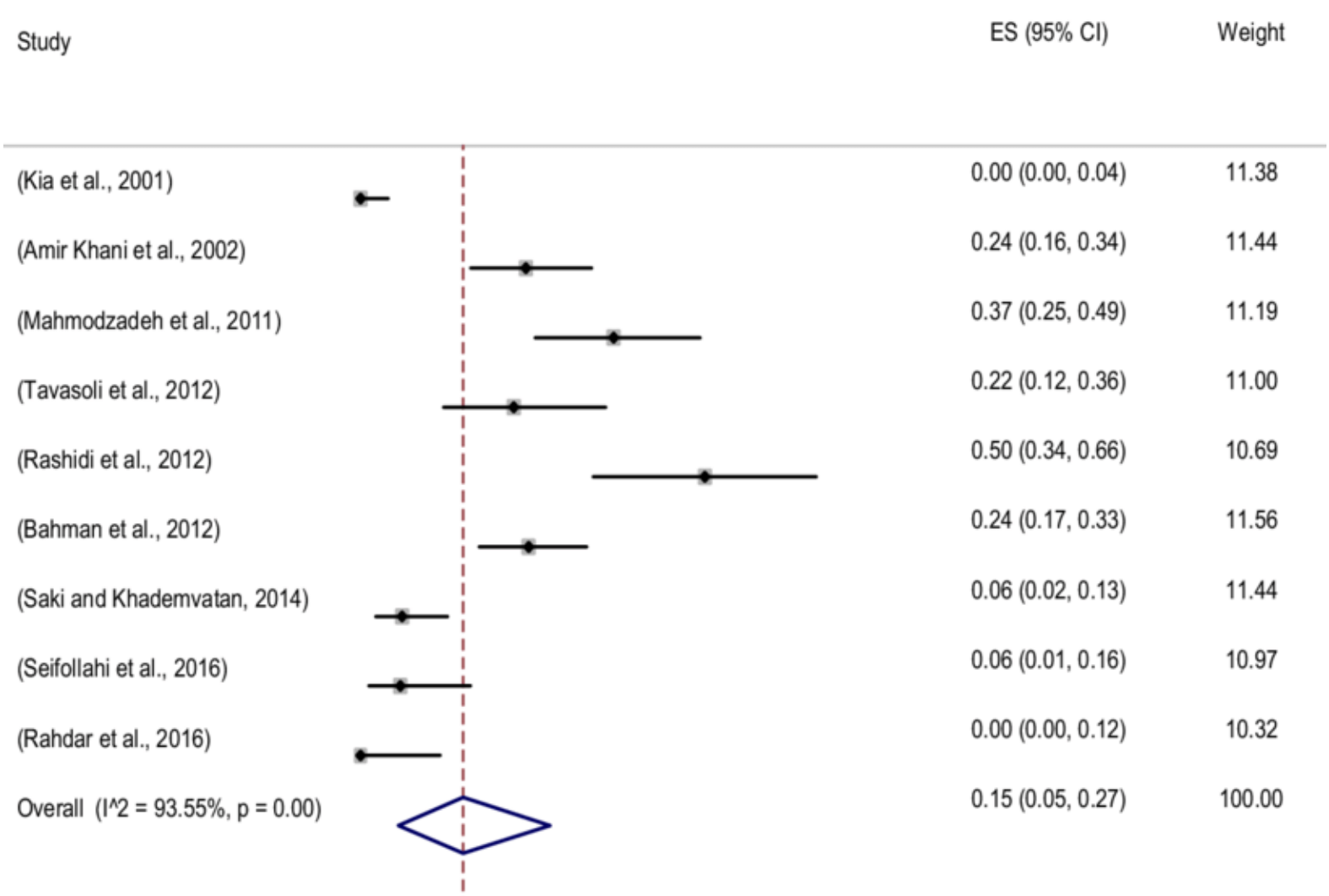
"An approach to the synthesis of findings from multiple studies that relies primarily on the use of words and text to summarize and explain the findings of the synthesis. It adopts a textual approach to the process of synthesis to 'tell the story' of the findings from the included studies." (Popay et al. 2006)

- Use tables to tabulate information extracted from included studies (e.g. population, number of included studies, study quality etc.)
- Tables should be structured to highlight similarity or differences of study outcomes
- Were the findings consistent (homogeneous) or inconsistent?

# Quantitative Synthesis (Meta-Analysis)

- Meta-analysis can be used to aggregate results or to pool data from different studies
- The outcome of a meta-analysis is an average effect size with an indication of how variable that effect size is between studies
- Meta-analysis involves three main steps:
  1. Decide which studies to be included in the meta-analysis
  2. Estimate an effect size for each individual study
  3. Combine the effect sizes from the individual studies to estimate and test the combined effect
- Results of the meta-analysis can be presented in a forest plot

# Contoh Forest Plot



# 3. Tahapan Reporting

1. Write Up the SLR Paper
2. Choose the Right Journal

# 1. Write Up the SLR Paper

1. Introduction
   - General introduction about the research
   - State the purpose of the review
   - Emphasize the reason(s) why the RQ is important
   - State the significance of the review work and how the project contributes to the body of knowledge of the field
2. Main Body
   1. Review method – briefly describe steps taken to conduct the review
   2. Results – findings from the review
   3. Discussion – implication of review for research & practice
3. Conclusions

# 2. Choose the Right Journal

- Some journals and conferences include a specific topic on SLR:
  - Information & Software Technology has an editor specializing in systematic reviews
  - Journal of Systems and Software
  - Expert Systems with Applications
  - IEEE Transactions on Software Engineering
  - Journal of Cleaner Production
  - European Journal of Internal Medicine
  - Journal of Supply Chain Management

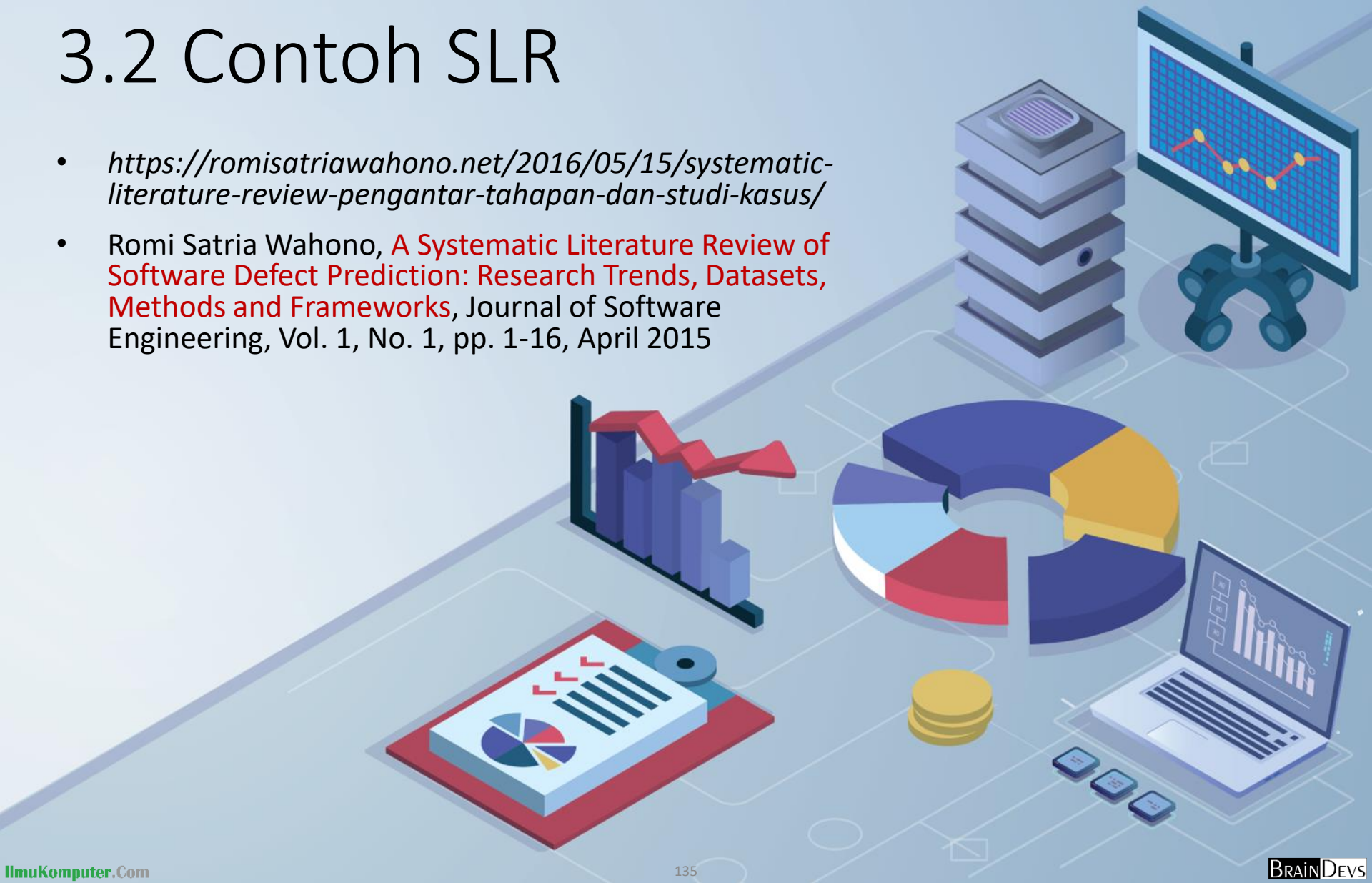
# Listing Jurnal Tujuan dan Nilai SJR/JIF

- Lakukan pendataan journal-journal yang ada di topik SLR yang kita tulis, urutkan berdasarkan rangking SJR atau JIF
- Publikasikan paper SLR kita ke journal yang sesuai dengan kualitas SLR yang kita lakukan
- A paper is an organized description of hypotheses, data and conclusions, intended to instruct the reader. If your research does not generate papers, it might just as well not have been done (Whitesides 2004)

| No | Journal Publications | SJR | Q Category |
|---|---|---|---|
| 1 | IEEE Transactions on Software Engineering | 3.39 | Q1 in Software |
| 2 | Information Sciences | 2.96 | Q1 in Information Systems |
| 3 | IEEE Transactions on Systems, Man, and Cybernetics | 2.76 | Q1 in Artificial Intelligence |
| 4 | IEEE Transactions on Knowledge and Data Engineering | 2.68 | Q1 in Information Systems |
| 5 | Empirical Software Engineering | 2.32 | Q1 in Software |
| 6 | Information and Software Technology | 1.95 | Q1 in Information Systems |
| 7 | Automated Software Engineering | 1.78 | Q1 in Software |
| 8 | IEEE Transactions on Reliability | 1.43 | Q1 in Software |
| 9 | Expert Systems with Applications | 1.36 | Q2 in Computer Science |
| 10 | Journal of Systems and Software | 1.09 | Q2 in Software |
| 11 | Software Quality Journal | 0.83 | Q2 in Software |
| 12 | IET Software | 0.55 | Q2 in Software |
| 13 | Advanced Science Letters | 0.24 | Q3 in Computer Science |
| 14 | Journal of Software | 0.23 | Q3 in Software |
| 15 | International Journal of Software Engineering and Its Application | 0.14 | Q4 in Software |

# 3.2 Contoh SLR

- *https://romisatriawahono.net/2016/05/15/systematic-literature-review-pengantar-tahapan-dan-studi-kasus/*
- Romi Satria Wahono, A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks, Journal of Software Engineering, Vol. 1, No. 1, pp. 1-16, April 2015



# Tahapan SLR

1. Formulate the review's research question
2. Develop the review's protocol

PLANNING

1. Identify the relevant literature
2. Perform selection of primary studies
3. Perform data extraction
4. Assess studies' quality
5. Conduct synthesis of evidence

CONDUCTING

1. Write up the SLR report/paper
2. Choose the Right Journal

REPORTING

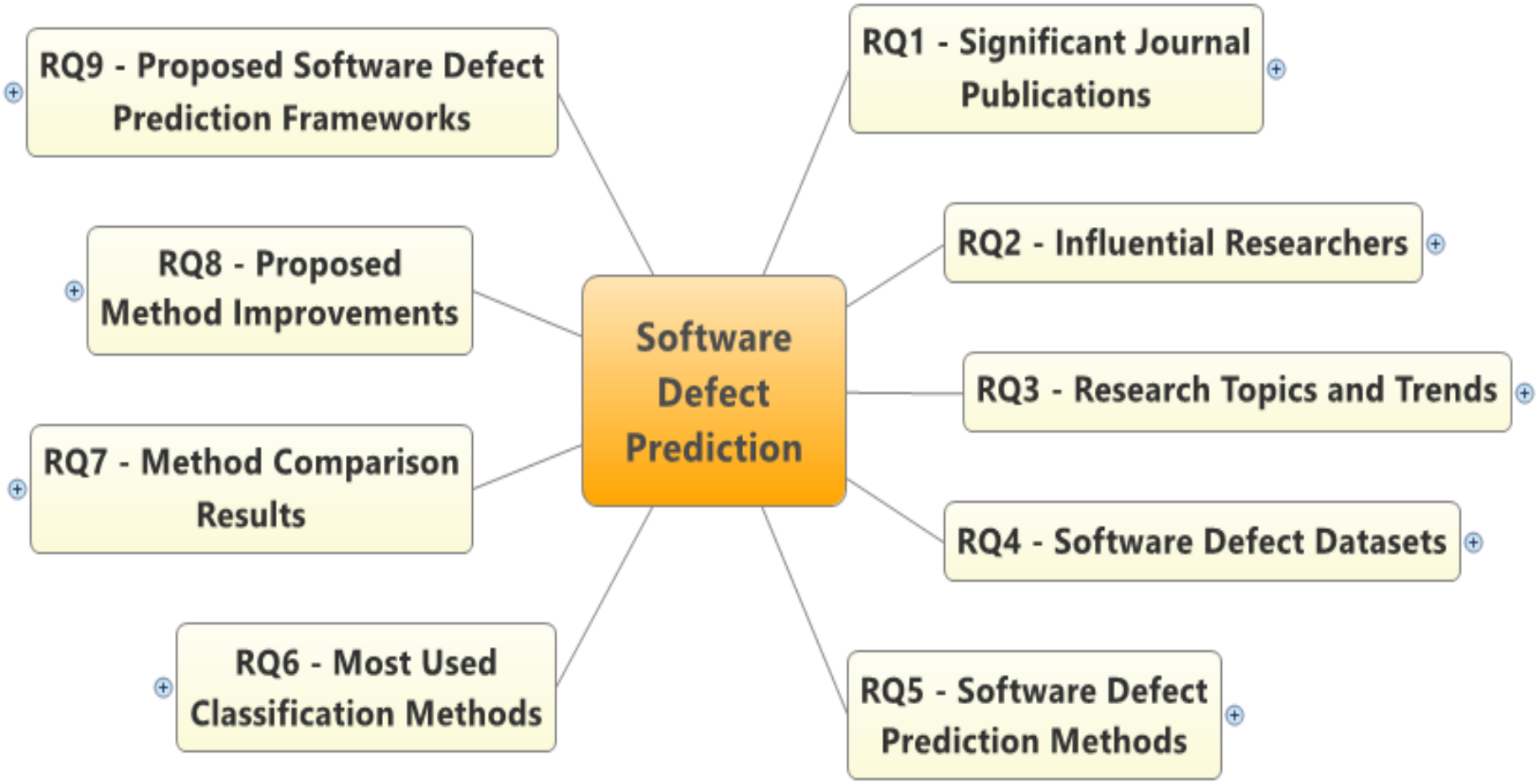
# 1. Contoh Planning

- *https://romisatriawahono.net/2016/05/15/systematic-literature-review-pengantar-tahapan-dan-studi-kasus/*
- Romi Satria Wahono, A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks, Journal of Software Engineering, Vol. 1, No. 1, pp. 1-16, April 2015

# Research Question (RQ)

| ID | Research Question |
|---|---|
| RQ1 | Which journal is the most significant software defect prediction journal? |
| RQ2 | Who are the most active and influential researchers in the software defect prediction field? |
| RQ3 | What kind of research topics are selected by researchers in the software defect prediction field? |
| RQ4 | What kind of datasets are the most used for software defect prediction? |
| RQ5 | What kind of methods are used for software defect prediction? |
| RQ6 | What kind of methods are used most often for software defect prediction? |
| RQ7 | Which method performs best when used for software defect prediction? |
| RQ8 | What kind of method improvements are proposed for software defect prediction? |
| RQ9 | What kind of frameworks are proposed for software defect prediction? |

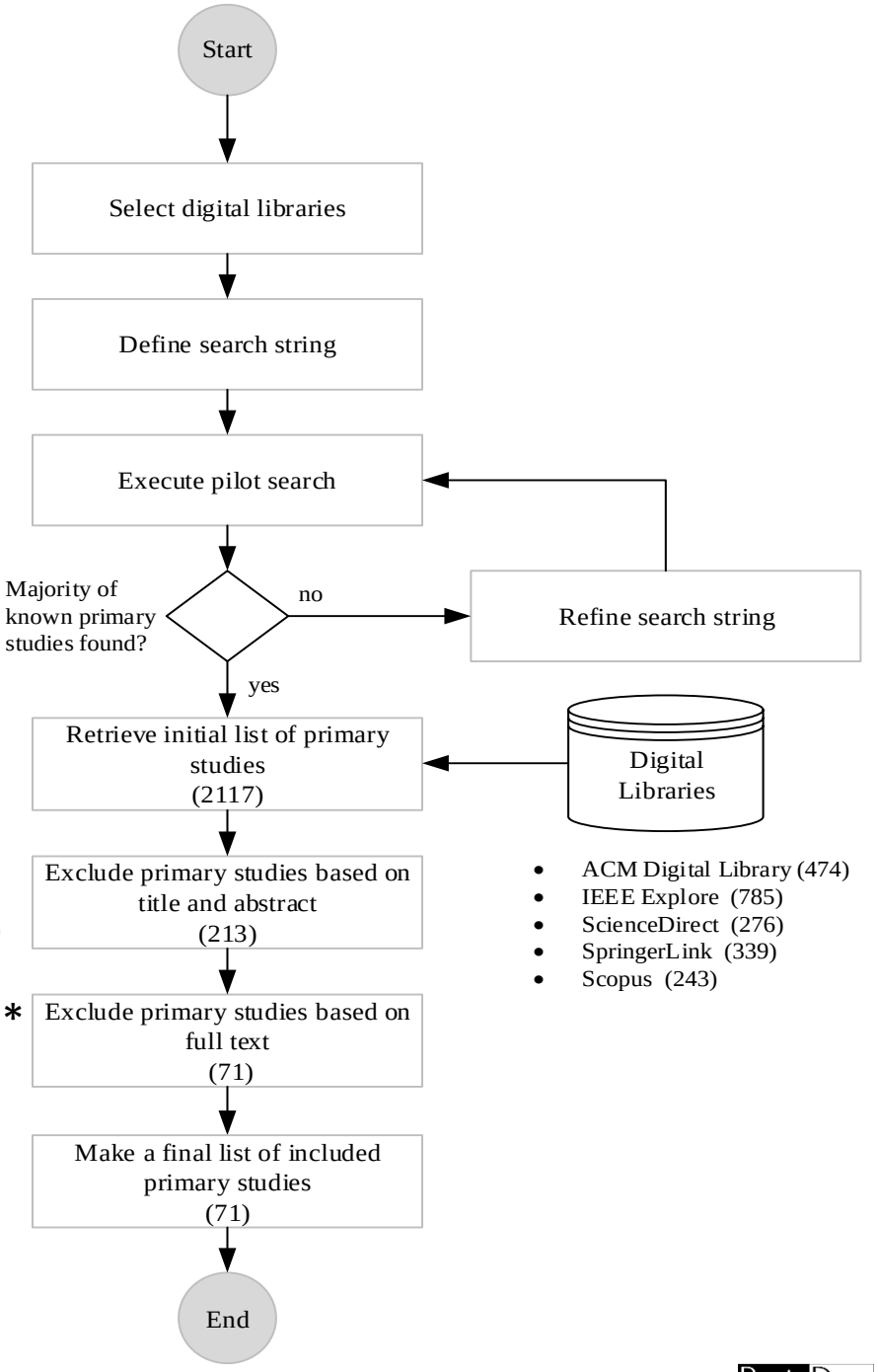
# Research Question (RQ)



# Studies Selection Strategy

- Publication Year:
  - ✓2000-2013
- Publication Type:
  - ✓Journal
  - ✓Conference Proceedings
- Search String:

  software
  AND
  (fault* OR defect* OR quality OR error-prone)
  AND
  (predict* OR prone* OR probability OR assess* OR detect* OR estimat* OR classificat*)
- Selected Studies:
  - ✓71



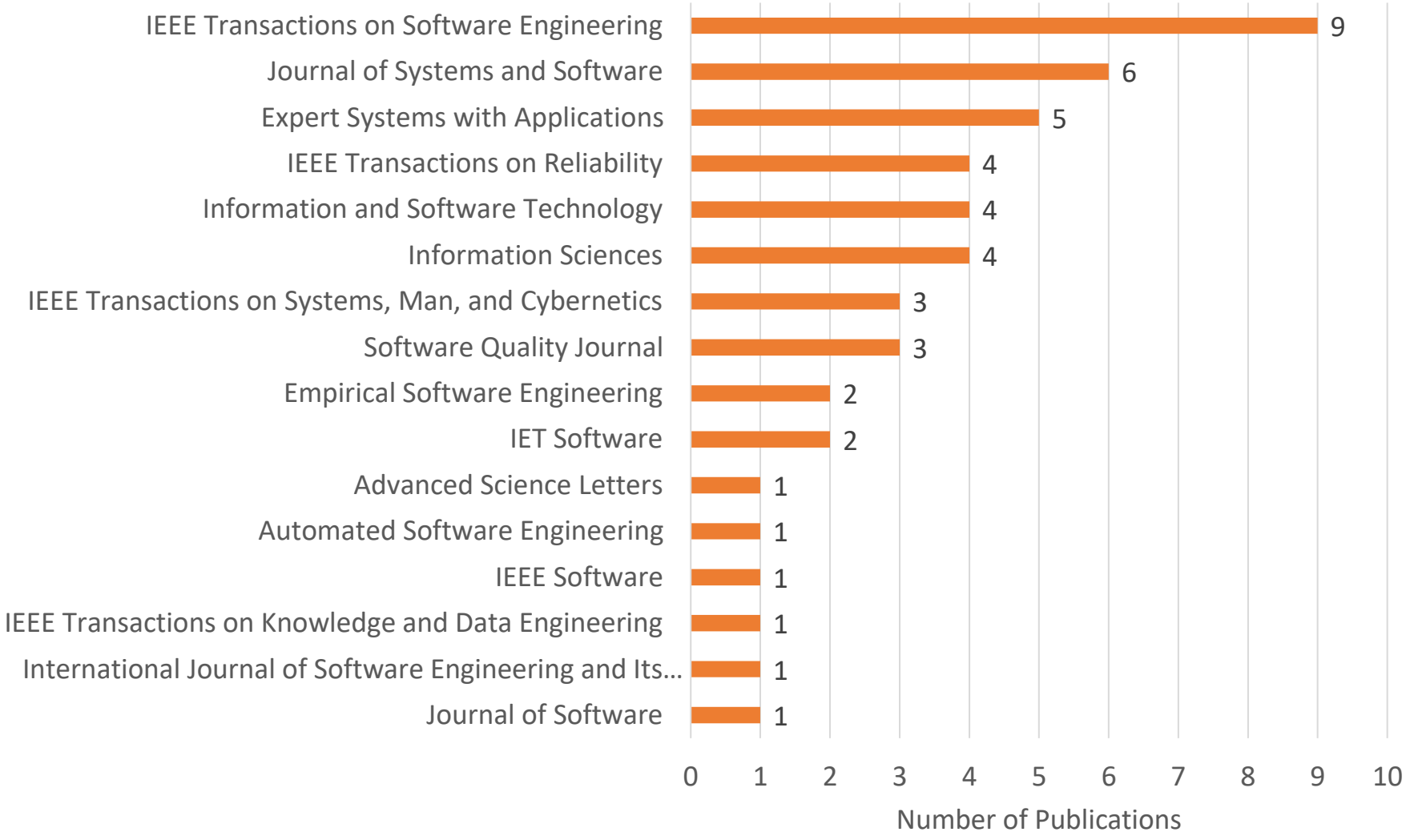
# Inclusion and Exclusion Criteria

| | |
|---|---|
| **Inclusion Criteria** | Studies in academic and industry using large and small scale data sets |
| | Studies discussing and comparing modeling performance in the area of software defect prediction |
| | For studies that have both the conference and journal versions, only the journal version will be included |
| | For duplicate publications of the same study, only the most complete and newest one will be included |
| **Exclusion Criteria** | Studies without a strong validation or including experimental results of software defect prediction |
| | Studies discussing defect prediction datasets, methods, frameworks in a context other than software defect prediction |
| | Studies not written in English |

# 2. Contoh Result

- *https://romisatriawahono.net/2016/05/15/systematic-literature-review-pengantar-tahapan-dan-studi-kasus/*
- Romi Satria Wahono, A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks, Journal of Software Engineering, Vol. 1, No. 1, pp. 1-16, April 2015

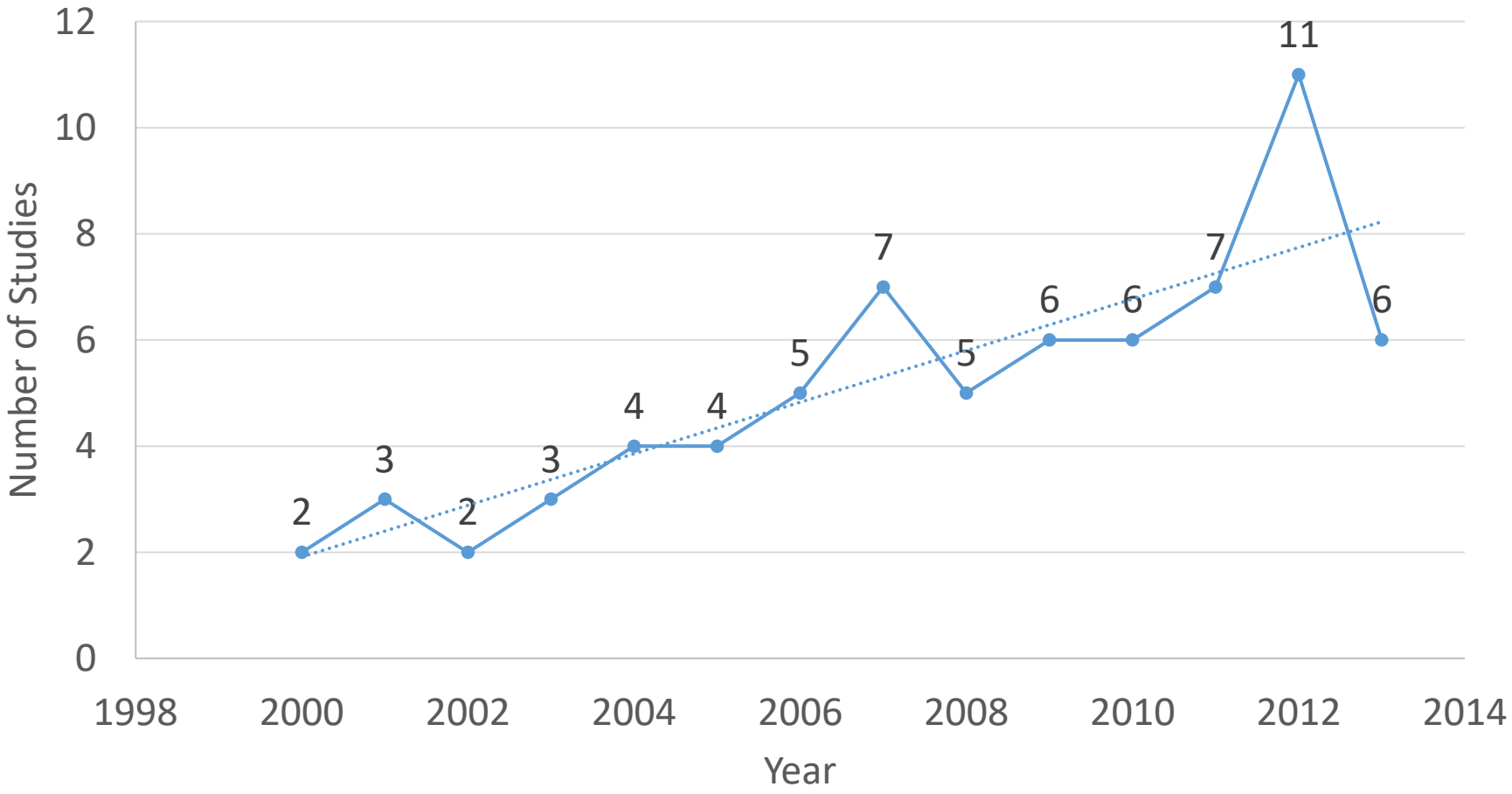
# RQ1: Significant Journal Publications



# Journal Quality Level of Selected Studies

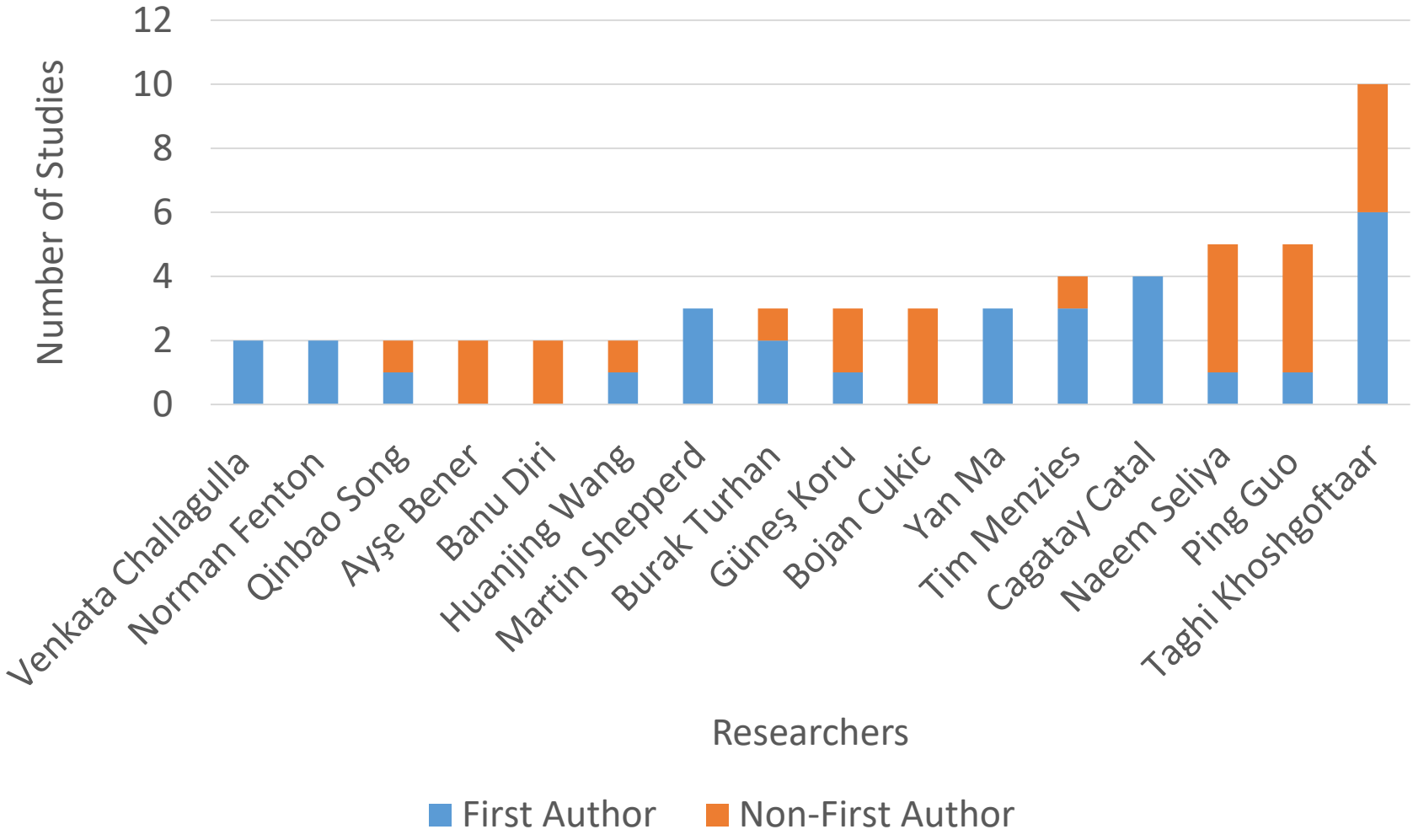
| No | Journal Publications | SJR | Q Category |
|---|---|---|---|
| 1 | IEEE Transactions on Software Engineering | 3.39 | Q1 in Software |
| 2 | Information Sciences | 2.96 | Q1 in Information Systems |
| 3 | IEEE Transactions on Systems, Man, and Cybernetics | 2.76 | Q1 in Artificial Intelligence |
| 4 | IEEE Transactions on Knowledge and Data Engineering | 2.68 | Q1 in Information Systems |
| 5 | Empirical Software Engineering | 2.32 | Q1 in Software |
| 6 | Information and Software Technology | 1.95 | Q1 in Information Systems |
| 7 | Automated Software Engineering | 1.78 | Q1 in Software |
| 8 | IEEE Transactions on Reliability | 1.43 | Q1 in Software |
| 9 | Expert Systems with Applications | 1.36 | Q2 in Computer Science |
| 10 | Journal of Systems and Software | 1.09 | Q2 in Software |
| 11 | Software Quality Journal | 0.83 | Q2 in Software |
| 12 | IET Software | 0.55 | Q2 in Software |
| 13 | Advanced Science Letters | 0.24 | Q3 in Computer Science |
| 14 | Journal of Software | 0.23 | Q3 in Software |
| 15 | International Journal of Software Engineering and Its Application | 0.14 | Q4 in Software |

# Distribution of Selected Studies by Year




- The interest in software defect prediction has changed over time
- Software defect prediction research is still very much relevant to this day

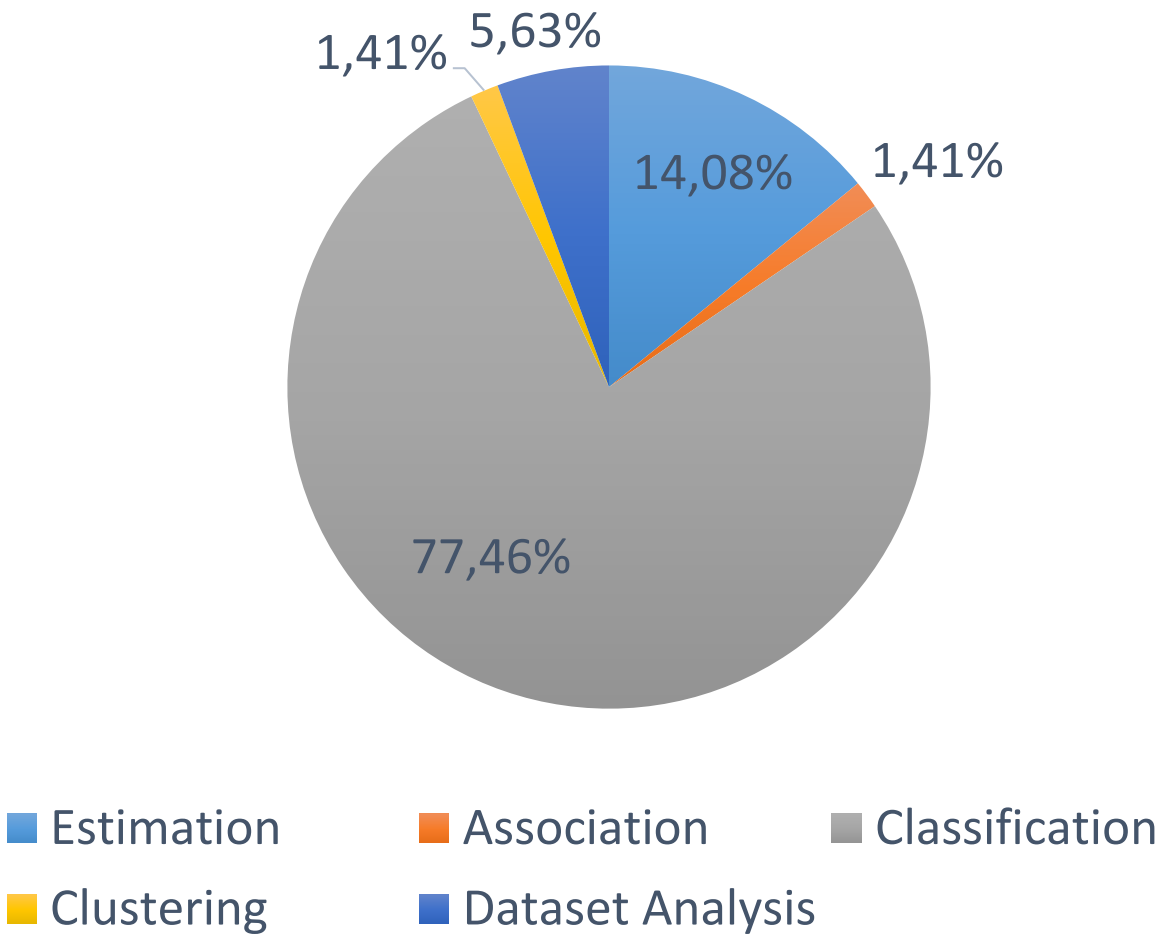
# RQ2: Influential Researchers



# RQ3: Research Topics and Trends

1. Estimating the number of defects remaining in software systems using estimation algorithm (Estimation)
2. Discovering defect associations using association rule algorithm (Association)
3. Classifying the defect-proneness of software modules, typically into two classes, defect-prone and not defect-prone, using classification algorithm (Classification)
4. Clustering the software defect based on object using clustering algorithm (Clustering)
5. Analyzing and pre-processing the software defect datasets (Dataset Analysis)

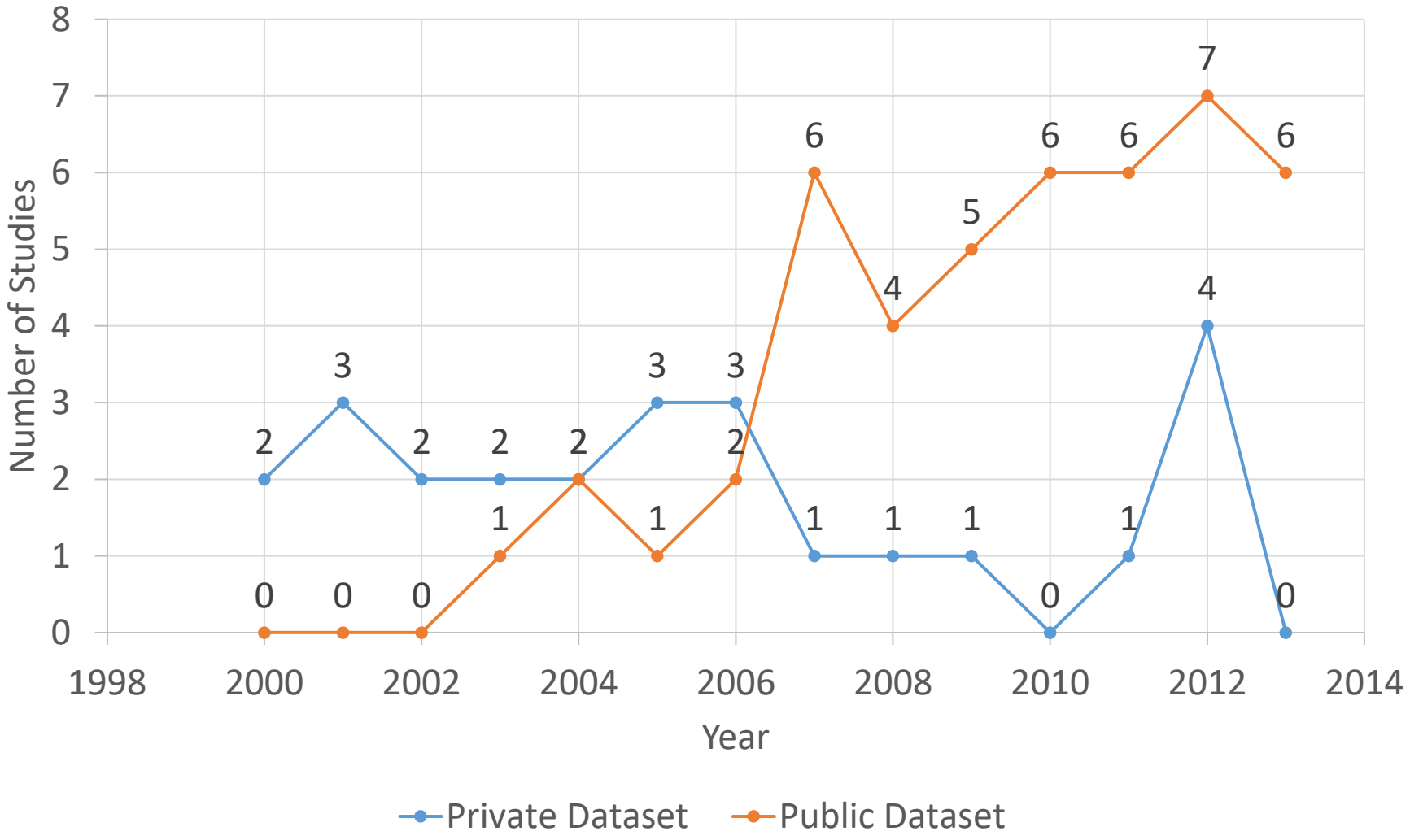
# Distribution of Research Topics and Trends



# Example Distribution of Research Topics and Trends

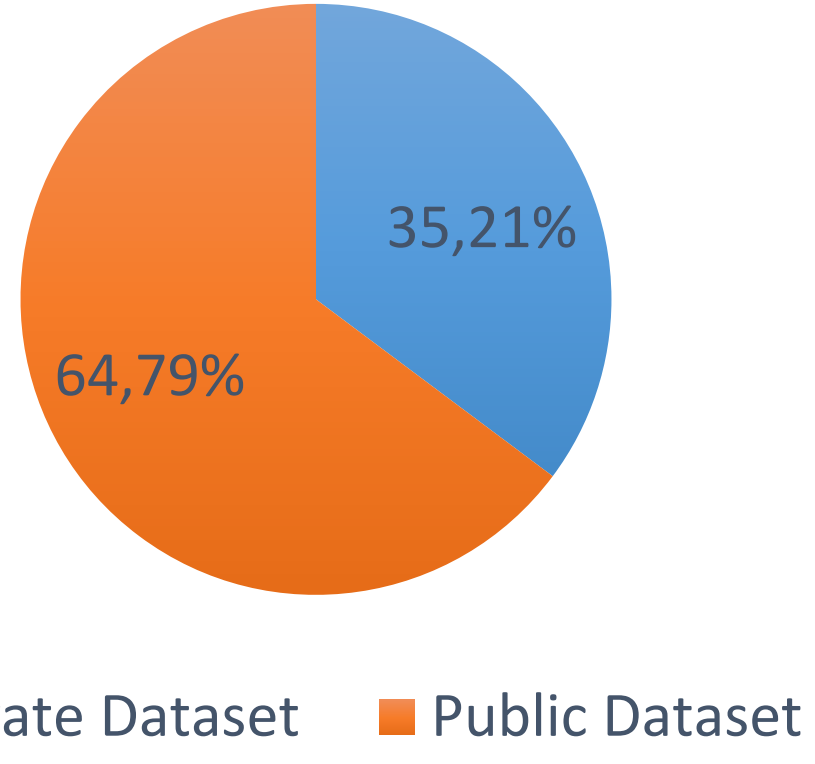
| Year | Primary Studies | Publications | Datasets | Topics |
|---|---|---|---|---|
| 2008 | (Lessmann et al., 2008) | IEEE Transactions on Software Engineering | Public | Classification |
| | (Bibi et al., 2008) | Expert Systems with Applications | Private | Estimation |
| | (Gondra, 2008) | Journal of Systems and Software | Public | Classification |
| | (Vandecruys et al., 2008) | Journal of Systems and Software | Public | Classification |
| | (Elish and Elish 2008) | Journal of Systems and Software | Public | Classification |
| 2012 | (Gray et al., 2012) | IET Software | Public | Dataset Analysis |
| | (Ying Ma, Luo, Zeng, & Chen, 2012) | Information and Software Technology | Public | Classification |
| | (Benaddy and Wakrim 2012) | International Journal of Software Engineering | Private | Estimation |
| | (Y. Peng, Wang, & Wang, 2012) | Information Sciences | Public | Classification |
| | (Zhang and Chang 2012) | International Conference on Natural Computation | Private | Estimation |
| | (Bishnu and Bhattacherjee 2012) | IEEE Transactions on Knowledge and Data Engineering | Private | Clustering |
| | (Sun, Song, & Zhu, 2012) | IEEE Transactions on Systems, Man, and Cybernetics | Public | Classification |
| | (Pelayo and Dick 2012) | IEEE Transactions on Reliability | Public | Classification |
| | (Jin, Jin, & Ye, 2012) | IET Software | Public | Classification |
| | (Cao, Qin, & Feng, 2012) | Advanced Science Letters | Public | Classification |
| 2013 | (Park et al., 2013) | Information Sciences | Public | Classification |
| | (Dejaeger, Verbraken, & Baesens, 2013) | IEEE Transactions on Software Engineering | Public | Classification |
| | (Shepperd, Song, Sun, & Mair, 2013) | IEEE Transactions on Software Engineering | Public | Dataset Analysis |
| | (Wang and Yao 2013) | IEEE Transactions on Reliability | Public | Classification |
| | (Peters, Menzies, Gong, & Zhang, 2013) | IEEE Transactions on Software Engineering | Public | Dataset Analysis |
| | (Radjenović et al., 2013) | Information and Software Technology | Public | Dataset Analysis |

# RQ4: Software Defect Datasets



# Distribution of Software Defect Datasets

- The use of public data sets makes the research repeatable, refutable, and verifiable (Catal & Diri 2009a)
- Since 2005 more public datasets were used
- NASA MDP repository have been developed in 2005 and researchers started to be aware regarding the use of public datasets

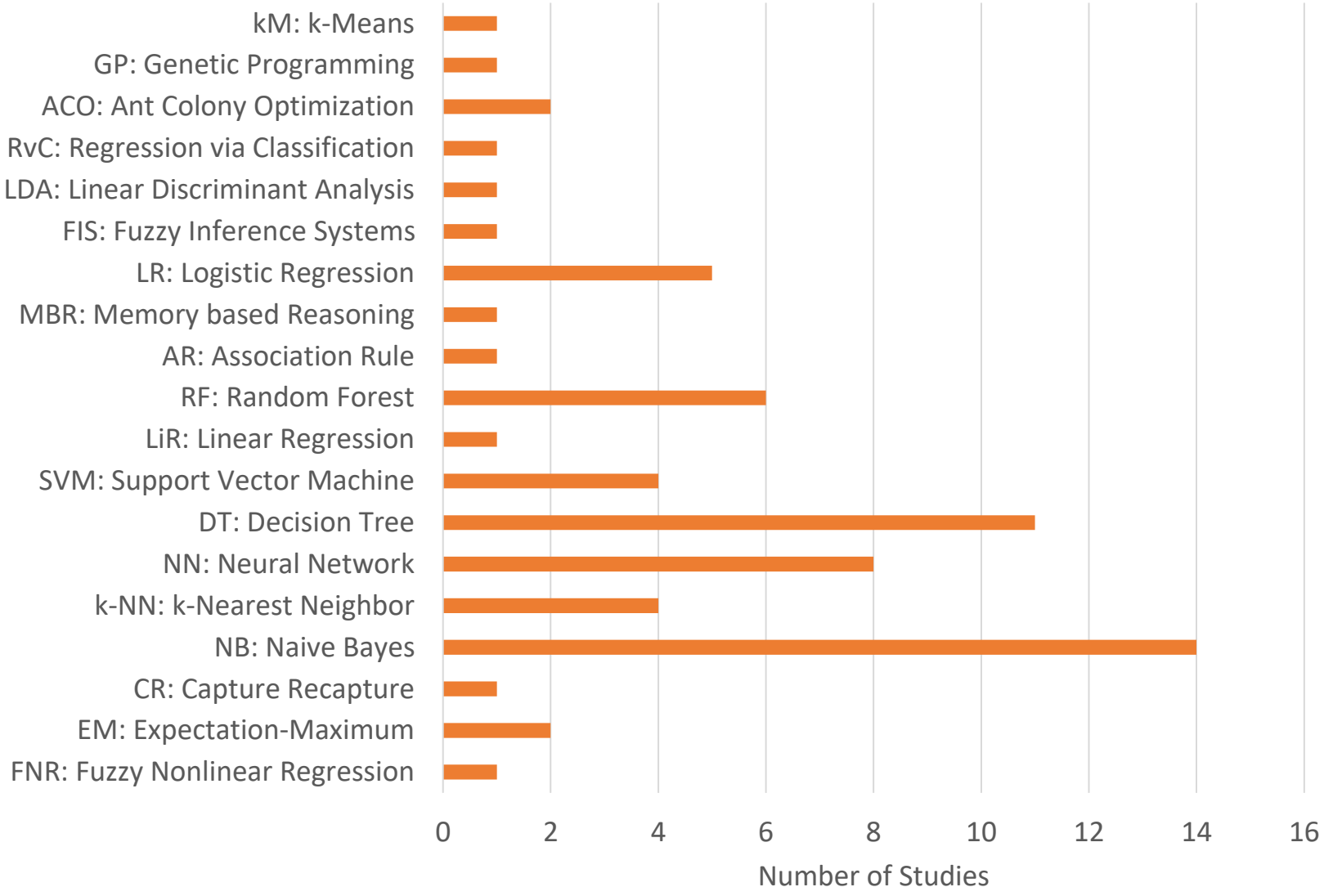


# NASA MDP Dataset

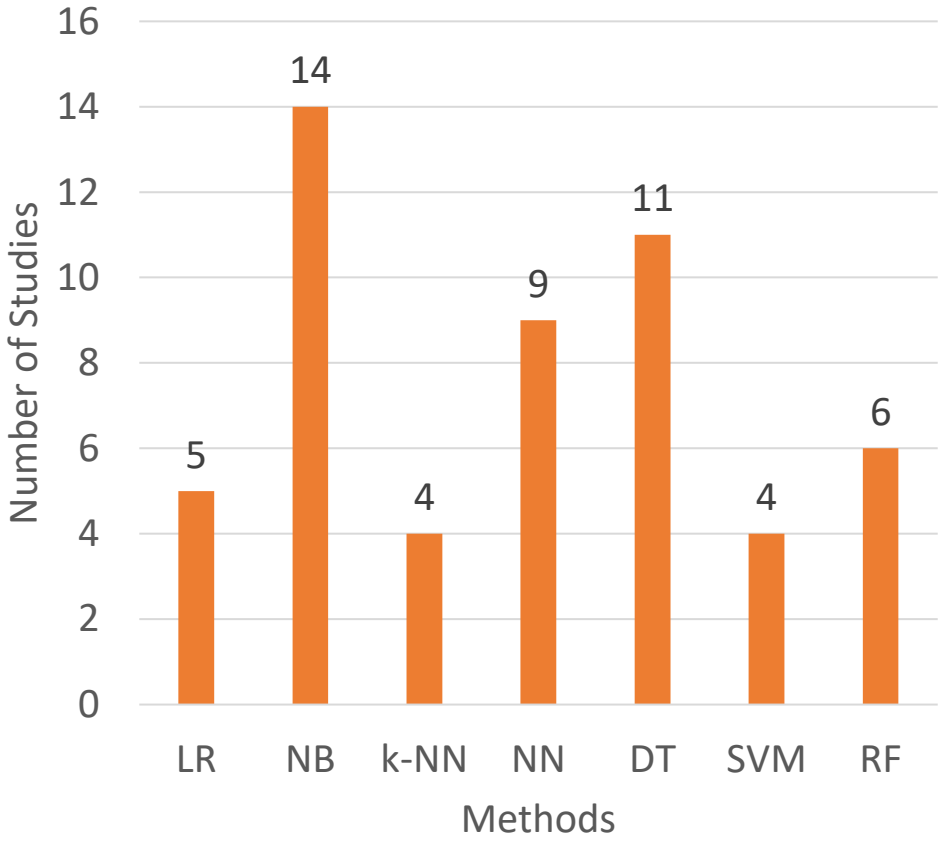
| Dataset | Project Description | Language | Number of Modules | Number of *fp* Modules | Faulty Percentage |
|---|---|---|---|---|---|
| CM1 | Spacecraft instrument | C | 505 | 48 | 12.21% |
| KC1 | Storage management for ground data | C++ | 1571 | 319 | 15.51% |
| KC3 | Storage management for ground data | Java | 458 | 42 | 18% |
| MC2 | Video guidance system | C | 127 | 44 | 34.65% |
| MW1 | Zero gravity experiment related to combustion | C | 403 | 31 | 10.23% |
| PC1 | Flight software from an earth orbiting satellite | C | 1059 | 76 | 8.04% |
| PC2 | Dynamic simulator for attitude control systems | C | 4505 | 23 | 1.01% |
| PC3 | Flight software for earth orbiting satellite | C | 1511 | 160 | 12.44% |
| PC4 | Flight software for earth orbiting satellite | C | 1347 | 178 | 12.72% |

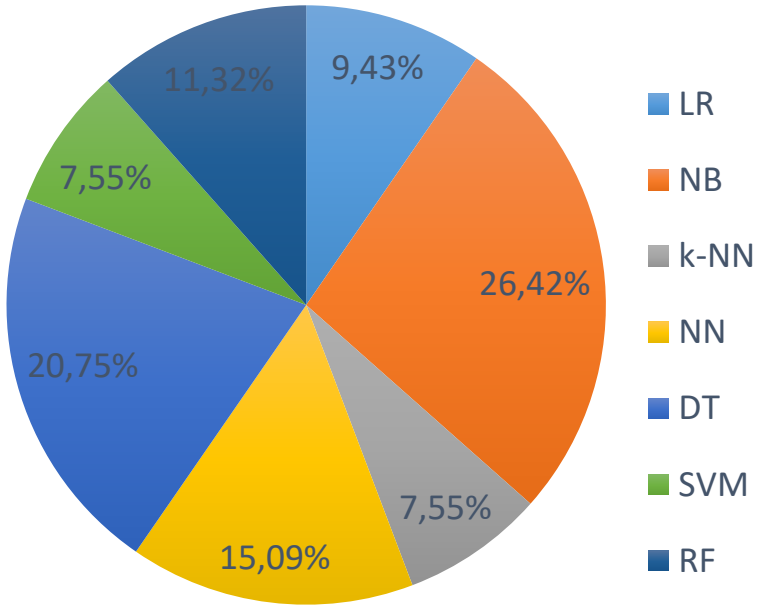
| Code Attributes | | Symbols | Description |
|---|---|---|---|
| LOC counts | LOC_total | | The total number of lines for a given module |
| | LOC_blank | | The number of blank lines in a module |
| | LOC_code_and_comment | NCSLOC | The number of lines which contain both code and comment in a module |
| | LOC_comments | | The number of lines of comments in a module |
| | LOC_executable | | The number of lines of executable code for a module |
| | number_of_lines | | Number of lines in a module |
| Halstead | content | $\mu$ | The halstead length content of a module $\mu = \mu_1 + \mu_2$ |
| | difficulty | D | The halstead difficulty metric of a module D = 1/L |
| | effort | E | The halstead effort metric of a module E = V/L |
| | error_est | B | The halstead error estimate metric of a module $B = E^{2/3}/1000$ |
| | length | N | The halstead length metric of a module $N = N_1+N_2$ |
| | level | L | The halstead level metric of a module $L = (2^* \mu_2)/ \mu_1^*N_2$ |
| | prog_time | T | The halstead programming time metric of a module T = E/18 |
| | volume | V | The halstead volume metric of a module $V = N^*\log_2(\mu_1+ \mu_2)$ |
| | num_operands | $N_1$ | The number of operands contained in a module |
| | num_operators | $N_2$ | The number of operators contained in a module |
| | num_unique_operands | $\mu_1$ | The number of unique operands contained in a module |
| | num_unique_operators | $\mu_2$ | The number of unique operators contained in a module |
| McCabe | cyclomatic_complexity | v(G) | The cyclomatic complexity of a module v(G) = e – n +2 |
| | cyclomatic_density | | v(G) / NCSLOC |
| | design_complexity | iv(G) | The design complexity of a module |
| | essential_complexity | ev(G) | The essential complexity of a module |
| Misc. | branch_count | | Branch count metrics |
| | call_pairs | | Number of calls to functions in a module |
| | condition_count | | Number of conditionals in a given module |
| | decision_count | | Number of decision points in a module |
| | decision_density | | condition_count / decision_count |
| | edge_count | | Number of edges found in a given module from one module to another |
| | essential_density | | Essential density is calculated as: (ev(G)-1)/(v(G)-1) |
| | parameter_count | | Number of parameters to a given module |
| | maintenance_severity | | Maintenance Severity is calculated as: ev(G)/v(G) |
| | modified_condition_count | | The effect of a condition affect a decision outcome by varying that condition only |
| | multiple_condition_count | | Number of multiple conditions within a module |
| | global_data_complexity | gdv(G) | the ratio of cyclomatic complexity of a module's structure to its parameter count |
| | global_data_density | | Global Data density is calculated as: gdv(G)/v(G) |
| | normalized_cyclo_cmplx | | v(G) / numbe_of_lines |
| | percent_comments | | Percentage of the code that is comments |
| | node_count | | Number of nodes found in a given module |

# RQ5: Software Defect Prediction Methods



# RQ6: Most Used Software Defect Prediction Methods



# RQ7: Method Comparison Results

- The comparisons and benchmarking result of the defect prediction using machine learning classifiers indicate that:
  - ✓Poor accuracy level is dominant (Lessmann et al. 2008)
  - ✓No significant performance differences could be detected (Lessmann et al. 2008)
  - ✓No particular classifiers that performs the best for all the data sets (Song et al. 2011) (Hall et al. 2012)

- The accurate and reliable classification algorithms to build a better prediction model is an open issue in software defect prediction

# RQ8: Method Improvement Efforts

- Researchers proposed some techniques for improving the accuracy of classifiers for software defect prediction
- Recent proposed techniques try to increase the prediction accuracy of a generated model:
  - ✓ By modifying and ensembling some machine learning methods (Mısırlı et al. 2011) (Tosun et al. 2008)
  - ✓ By using boosting algorithm (Zheng 2010) (Jiang et al. 2011)
  - ✓ by adding feature selection (Gayatri et al. 2010) (Khoshgoftaar & Gao, 2009) (Song et al. 2011)
  - ✓ By using parameter selection for some classifiers (Peng & Wang 2010) (Lin et al. 2008) (Guo et al. 2008)
- While considerable works have been done separately, limited research can be found on investigating them all together

# RQ9: Existing Frameworks

Three frameworks have been highly cited and influential in software defect prediction field

**Menzies** Framework

(Menzies et al. 2007)

**Lessmann** Framework

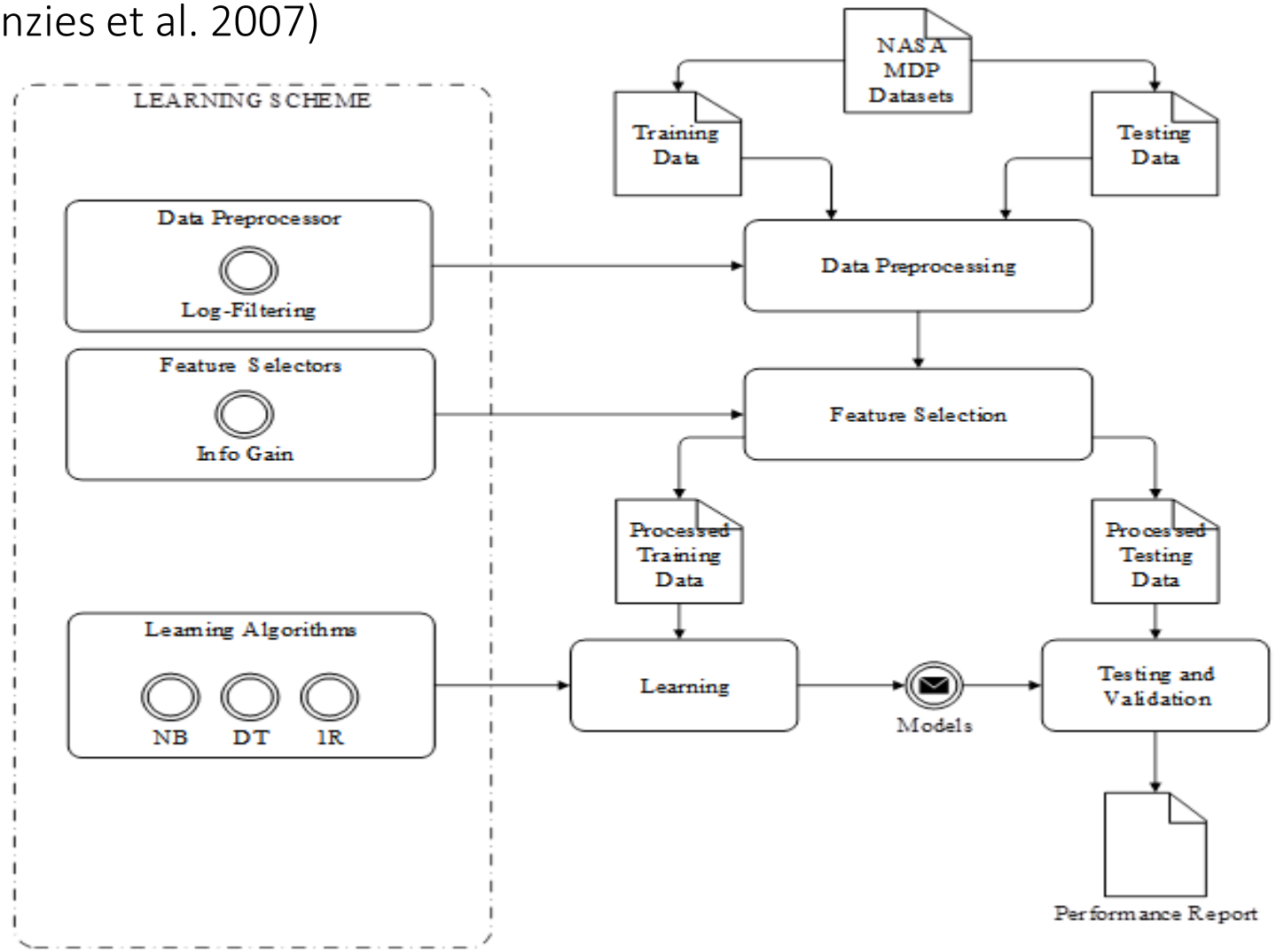
(Lessmann et al. 2008)

**Song** Framework

(Song et al. 2011)

# Menzies Framework

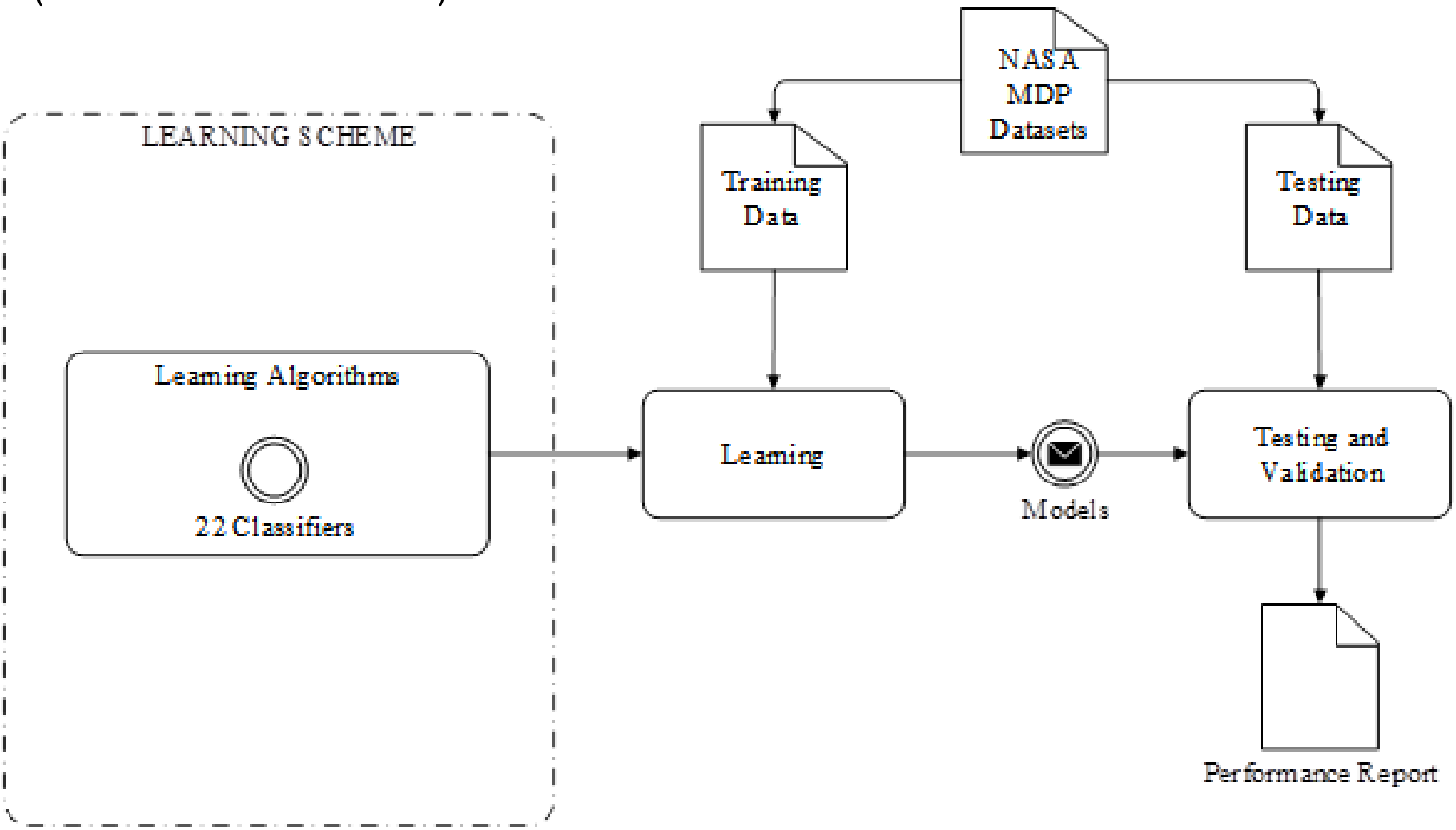
(Menzies et al. 2007)




| Framework | Dataset | Data Preprocessor | Feature Selectors | Meta-learning | Classifiers | Parameter Selectors | Validation Methods | Evaluation Methods |
|---|---|---|---|---|---|---|---|---|
| (Menzies et al. 2007) | NASA MDP | **Log Filtering** | **Info Gain** | - | **3 algorithms (DT, 1R, NB)** | - | 10-Fold X Validation | ROC Curve (AUC) |

# Lessmann Framework

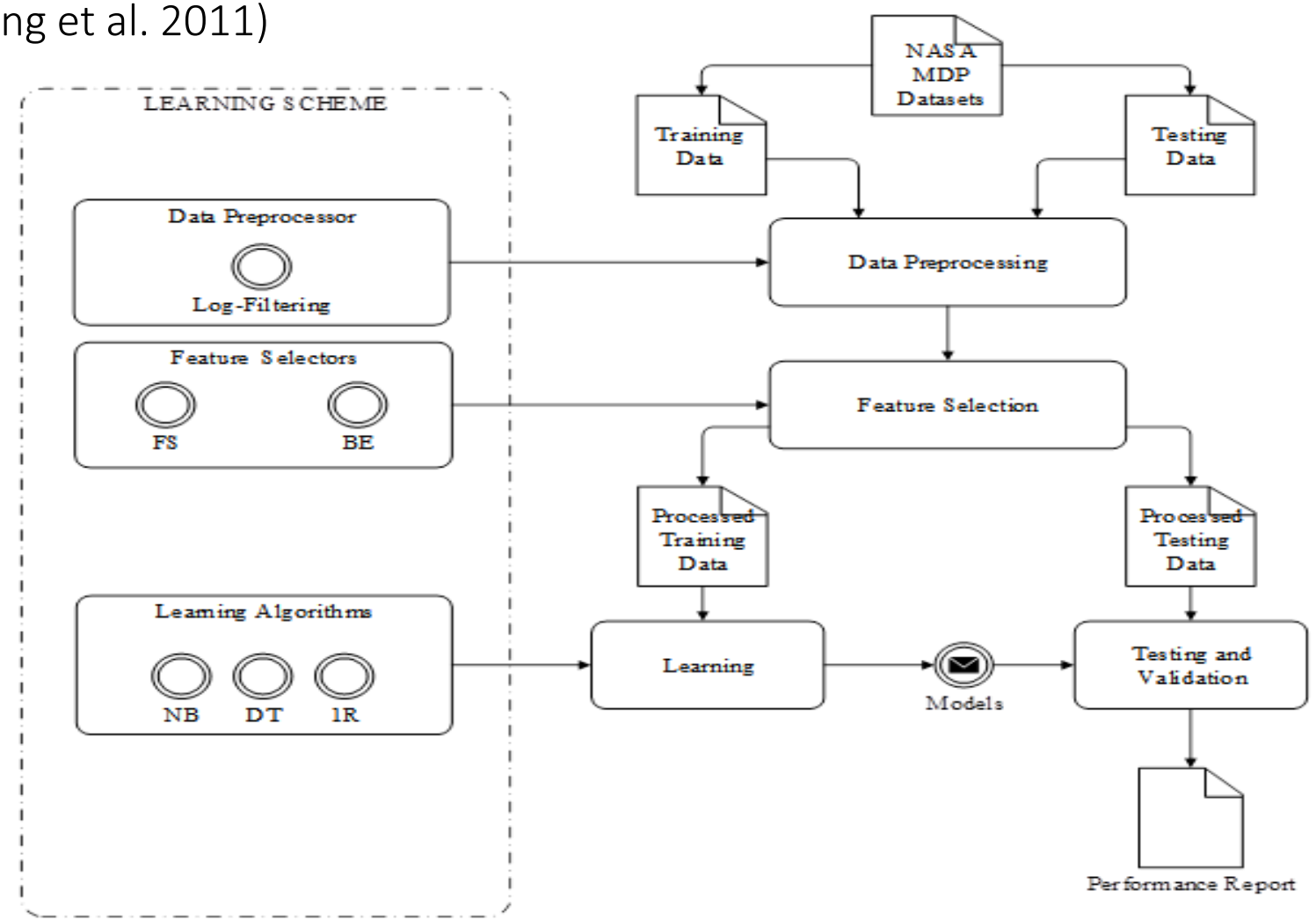
(Lessmann et al. 2008)




| Framework | Dataset | Data Preprocessor | Feature Selectors | Meta-learning | Classifiers | Parameter Selectors | Validation Methods | Evaluation Methods |
|---|---|---|---|---|---|---|---|---|
| (Lessman et al. 2008) | NASA MDP | - | - | - | **22 algorithms** | - | 10-Fold X Validation | ROC Curve (AUC) |

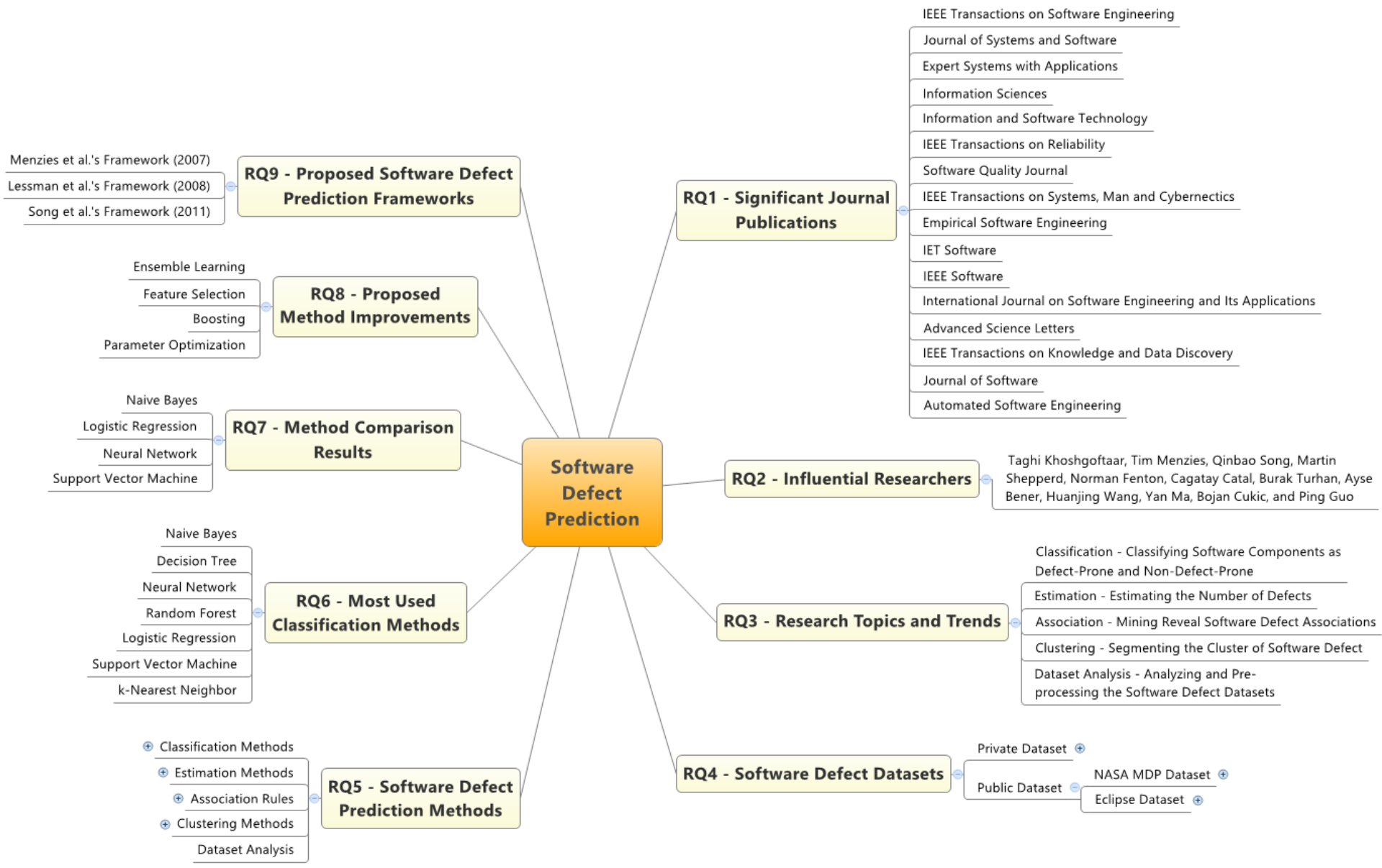
# Song Framework

(Song et al. 2011)




| Framework | Dataset | Data Preprocessor | Feature Selectors | Meta-learning | Classifiers | Parameter Selectors | Validation Methods | Evaluation Methods |
|---|---|---|---|---|---|---|---|---|
| (Song et al. 2011) | NASA MDP | **Log Filtering** | **FS, BE** | - | **3 algorithms (DT, 1R, NB)** | - | 10-Fold X Validation | ROC Curve (AUC) |

# Mind Map of the SLR Results



# SLR Melahirkan Research Gaps

Dari Hasil SLR, Kita Menemukan Research Gaps yang Menjadi Kandidat Masalah Penelitian yang Kita Angkat pada Penelitian Kita

# Gap Analysis of Framework

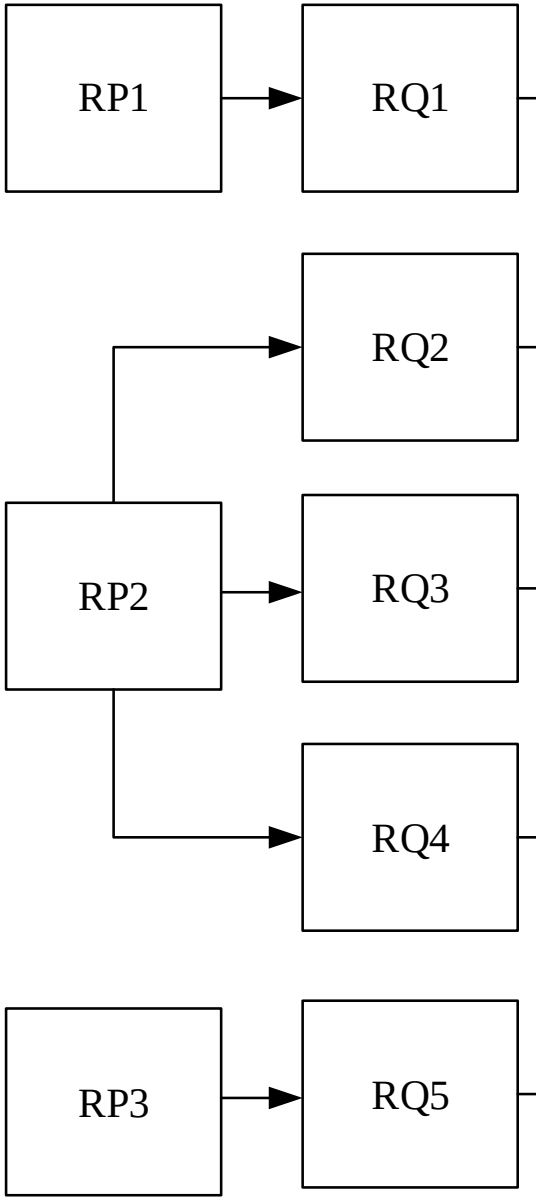
- The comparisons and benchmarking result of the defect prediction using machine learning classifiers indicate that:
  - ✓Poor accuracy level is dominant (Lessmann et al. 2008)
  - ✓No significant performance differences could be detected (Lessmann et al. 2008)
  - ✓No particular classifiers that performs the best for all the data sets (Song et al. 2011) (Hall et al. 2012)
- Noisy attribute predictors and imbalanced class distribution of software defect datasets result in inaccuracy of classification models
- Neural network and support vector machine have strong fault tolerance and strong ability of nonlinear dynamic processing of software fault data, but practicability of neural network and support vector machine are limited due to difficulty of selecting appropriate parameters

# Research Problems (RP)

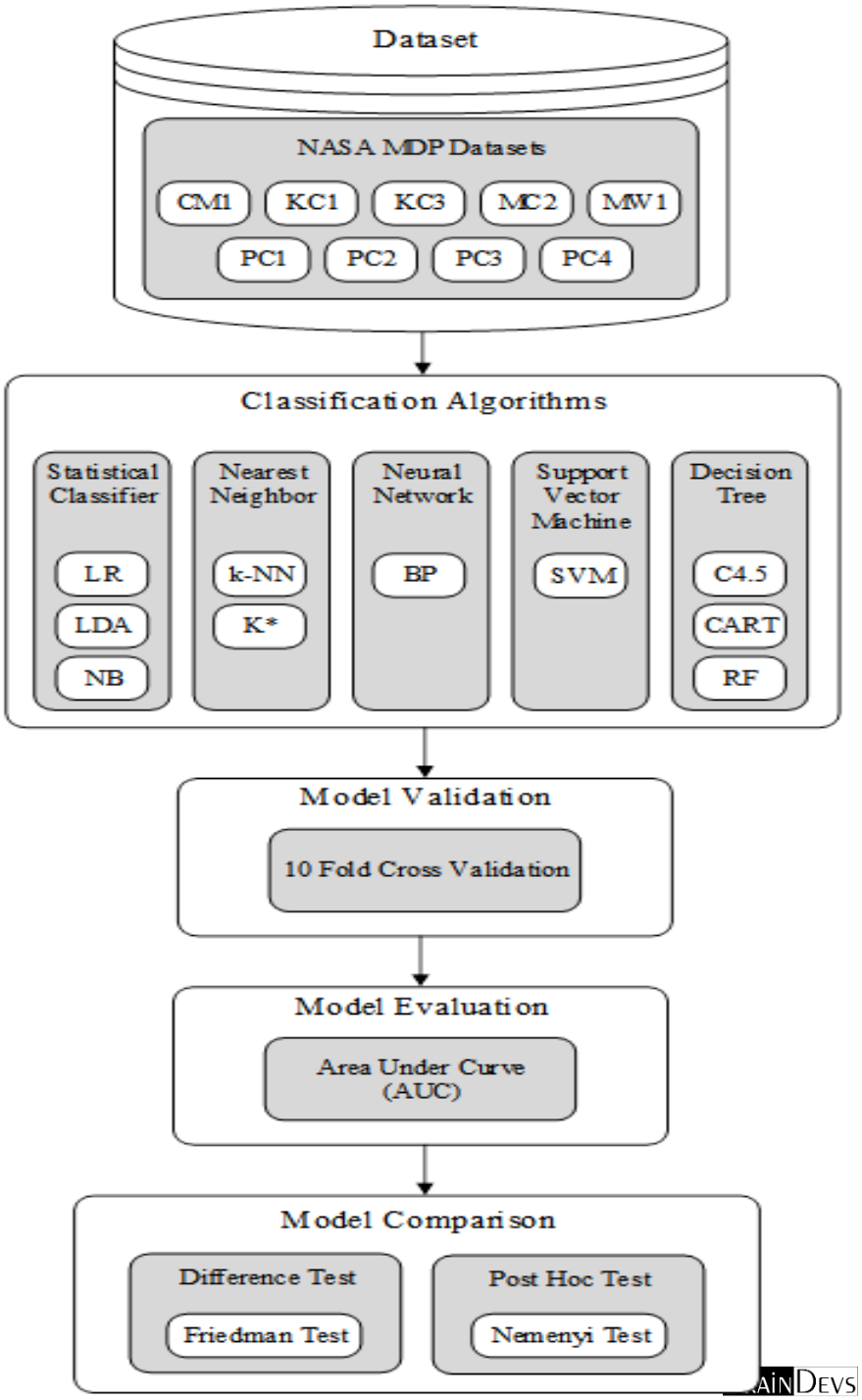
**RP1** While many studies on software defect prediction report the comparative performance of the classification algorithms used, but there is no strong consensus on which classifiers perform best when individual studies are looked separately

**RP2** Noisy attribute predictors and imbalanced class distribution of software defect datasets result in inaccuracy of classification models

**RP3** Neural network has strong fault tolerance and strong ability of nonlinear dynamic processing of software fault data, but practicability of neural network is limited due to difficulty of selecting appropriate parameters



# A Comparison Framework of Classification Models for Software Defect Prediction (CF SDP)



# AUC and Friedman Test Results

| | CM1 | KC1 | KC3 | MC2 | MW1 | PC1 | PC2 | PC3 | PC4 | *M* | *R* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| LR | **0.763** | **0.801** | **0.713** | **0.766** | 0.726 | **0.852** | 0.849 | **0.81** | 0.894 | **0.797** | **1.44** |
| LDA | 0.471 | 0.536 | 0.447 | 0.503 | 0.58 | 0.454 | 0.577 | 0.524 | 0.61 | 0.522 | 8.33 |
| NB | 0.734 | 0.786 | 0.67 | 0.739 | 0.732 | 0.781 | 0.811 | 0.756 | 0.838 | 0.761 | 3 |
| k-NN | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 8.778 |
| K* | 0.6 | 0.678 | 0.562 | 0.585 | 0.63 | 0.652 | 0.754 | 0.697 | 0.76 | 0.658 | 5.33 |
| BP | 0.713 | 0.791 | 0.647 | 0.71 | 0.625 | 0.784 | 0.918 | 0.79 | 0.883 | 0.762 | 3.22 |
| SVM | 0.753 | 0.752 | 0.642 | 0.761 | 0.714 | 0.79 | 0.534 | 0.75 | **0.899** | 0.733 | 3.33 |
| C4.5 | 0.565 | 0.515 | 0.497 | 0.455 | 0.543 | 0.601 | 0.493 | 0.715 | 0.723 | 0.567 | 7.78 |
| CART | 0.604 | 0.648 | 0.637 | 0.482 | 0.656 | 0.574 | 0.491 | 0.68 | 0.623 | 0.599 | 6.89 |
| RF | 0.573 | 0.485 | 0.477 | 0.525 | **0.74** | 0.618 | 0.649 | 0.678 | 0.2 | 0.549 | 6.89 |

- LR is dominant in most datasets
- *R* rank: LR has the highest rank, followed by NB, BP, and SVM
- *M* results: no excellent or good models, and a few fair models

| AUC | Meaning | Symbol |
|---|---|---|
| 0.90 - 1.00 | excellent classification | |
| 0.80 - 0.90 | good classification | |
| 0.70 - 0.80 | fair classification | |
| 0.60 - 0.70 | poor classification | |
| < 0.60 | failure | |

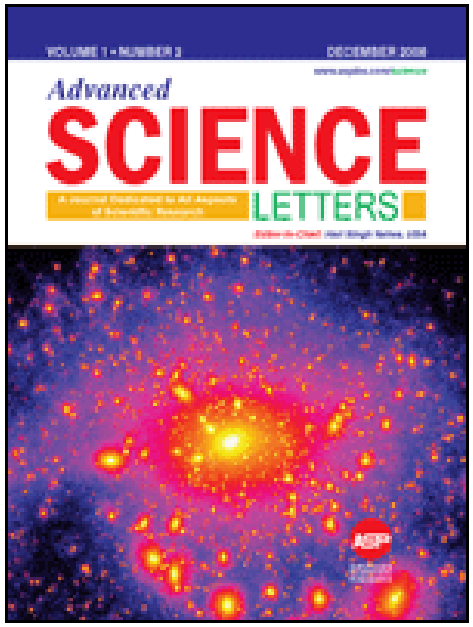
# *P-value* of Nemenyi Post Hoc Test

| | LR | LDA | NB | k-NN | K* | BP | SVM | C4.5 | CART | RF |
|---|---|---|---|---|---|---|---|---|---|---|
| LR | 1 | **0.0001** | 0.986 | **0.0001** | 0.164 | 0.965 | 0.949 | **0.000** | **0.005** | **0.005** |
| LDA | **0.0001** | 1 | **0.007** | 1.000 | 0.526 | **0.013** | **0.017** | 1.000 | 0.992 | 0.992 |
| NB | 0.986 | **0.007** | 1 | **0.002** | 0.831 | 1.000 | 1.000 | **0.028** | 0.164 | 0.164 |
| k-NN | **0.0001** | 1.000 | **0.002** | 1 | 0.318 | **0.004** | **0.005** | 1.000 | 0.949 | 0.949 |
| K* | 0.164 | 0.526 | 0.831 | 0.318 | 1 | 0.901 | 0.927 | 0.789 | 0.986 | 0.986 |
| BP | 0.965 | **0.013** | 1.000 | **0.004** | 0.901 | 1 | 1.000 | **0.046** | 0.232 | 0.232 |
| SVM | 0.949 | **0.017** | 1.000 | **0.005** | 0.927 | 1.000 | 1 | 0.058 | 0.273 | 0.273 |
| C4.5 | **0.000** | 1.000 | **0.028** | 1.000 | 0.789 | **0.046** | 0.058 | 1 | 1.000 | 1.000 |
| CART | **0.005** | 0.992 | 0.164 | 0.949 | 0.986 | 0.232 | 0.273 | 1.000 | 1 | 1.000 |
| RF | **0.005** | 0.992 | 0.164 | 0.949 | 0.986 | 0.232 | 0.273 | 1.000 | 1.000 | 1 |

- If *P value < 0.05* (boldfaced print), it indicate that there is significant different between two classifiers
- Based on significant difference results, there is no significant difference between LR, NB, BP, and SVM models

# Research Publication on RQ1

Romi Satria Wahono, Nanna Suryana Herman and Sabrina Ahmad, A Comparison Framework of Classification Models for Software Defect Prediction, Advanced Science Letters, Vol. 20, No. 8, August 2014



# Research Result on RQ3

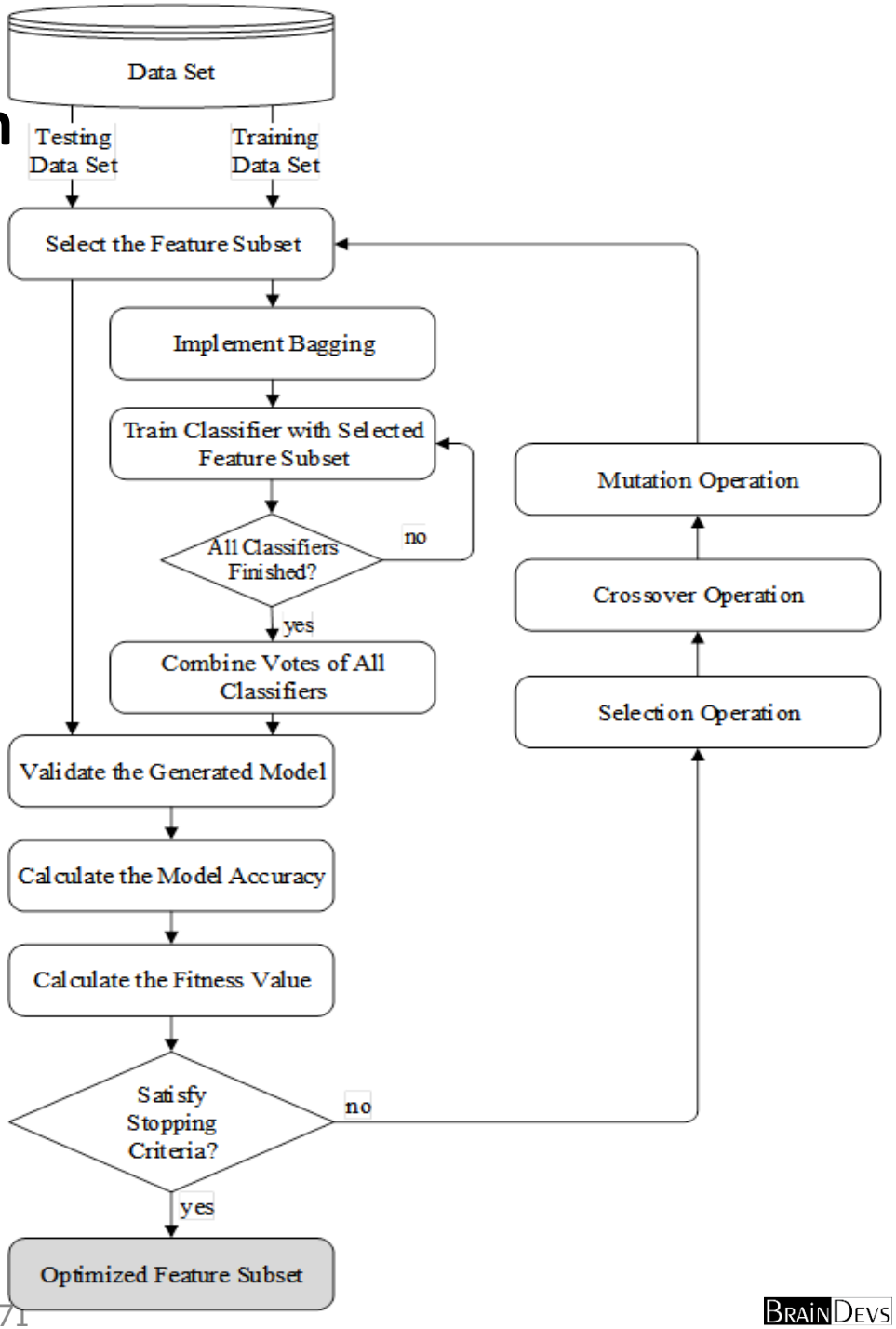
| Research Problems (RP) | | Research Questions (RQ) | | Research Objectives (RO) | |
|---|---|---|---|---|---|
| RP2 | **Noisy attribute predictors and imbalanced class distribution of software defect datasets result in inaccuracy of classification models** | RQ2 | How does the integration between genetic algorithm based feature selection and bagging technique affect the accuracy of software defect prediction? | RO2 | To develop a hybrid genetic algorithm based feature selection and bagging technique for improving the accuracy of software defect prediction |
| | | **RQ3** | **How does the integration between particle swarm optimization based feature selection and bagging technique affect the accuracy of software defect prediction?** | **RO3** | **To develop a hybrid particle swarm optimization based feature selection and bagging technique for improving the accuracy of software defect prediction** |
| | | RQ4 | Which metaheuristic optimization techniques perform best when used in feature selection of software defect prediction? | RO4 | To identify the best metaheuristic optimization techniques when used in feature selection of software defect prediction |

# A Hybrid Genetic Algorithm based Feature Selection and Bagging Technique (GAFS+B)

- Every chromosome is evaluated by the fitness function Equation

$$fitness = W_A \times A + W_F \times \left( P + \left( \sum_{i=1}^{n_f} C_i \times F_i \right) \right)^{-1}$$

- Where
  - $A$: classification accuracy
  - $F_i$: feature value
  - $W_A$: weight of classification accuracy
  - $W_F$: feature weight
  - $C_i$: feature cost
- When ending condition is satisfied, the operation ends, otherwise, continue with the next genetic operation



# Results: Without GAFS+B

| Classifiers | | CM1 | KC1 | KC3 | MC2 | MW1 | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|---|---|---|---|---|---|
| Statistical Classifier | LR | 0.763 | 0.801 | 0.713 | 0.766 | 0.726 | 0.852 | 0.849 | 0.81 | 0.894 |
| | LDA | 0.471 | 0.536 | 0.447 | 0.503 | 0.58 | 0.454 | 0.577 | 0.524 | 0.61 |
| | NB | 0.734 | 0.786 | 0.67 | 0.739 | 0.732 | 0.781 | 0.811 | 0.756 | 0.838 |
| Nearest Neighbor | k-NN | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | K* | 0.6 | 0.678 | 0.562 | 0.585 | 0.63 | 0.652 | 0.754 | 0.697 | 0.76 |
| Neural Network | BP | 0.713 | 0.791 | 0.647 | 0.71 | 0.625 | 0.784 | 0.918 | 0.79 | 0.883 |
| Support Vector Machine | SVM | 0.753 | 0.752 | 0.642 | 0.761 | 0.714 | 0.79 | 0.534 | 0.75 | 0.899 |
| Decision Tree | C4.5 | 0.565 | 0.515 | 0.497 | 0.455 | 0.543 | 0.601 | 0.493 | 0.715 | 0.723 |
| | CART | 0.604 | 0.648 | 0.637 | 0.482 | 0.656 | 0.574 | 0.491 | 0.68 | 0.623 |
| | RF | 0.573 | 0.485 | 0.477 | 0.525 | 0.74 | 0.618 | 0.649 | 0.678 | 0.2 |

# Results: With GAFS+B

| Classifiers | | CM1 | KC1 | KC3 | MC2 | MW1 | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|---|---|---|---|---|---|
| Statistical Classifier | LR | 0.753 | 0.795 | 0.691 | 0.761 | **0.742** | 0.852 | 0.822 | 0.813 | **0.901** |
| | LDA | **0.592** | **0.627** | **0.635** | **0.64** | **0.674** | **0.637** | **0.607** | **0.635** | **0.715** |
| | NB | 0.702 | **0.79** | **0.677** | 0.739 | 0.724 | **0.799** | 0.805 | **0.78** | **0.861** |
| Nearest Neighbor | k-NN | **0.666** | **0.689** | **0.67** | **0.783** | **0.656** | **0.734** | **0.554** | **0.649** | **0.732** |
| | K* | **0.71** | **0.822** | 0.503 | **0.718** | **0.68** | **0.876** | **0.877** | **0.816** | **0.893** |
| Neural Network | BP | **0.744** | **0.797** | **0.707** | **0.835** | **0.689** | **0.829** | 0.905 | **0.799** | **0.921** |
| Support Vector Machine | SVM | 0.667 | **0.767** | 0.572 | **0.747** | **0.659** | 0.774 | 0.139 | 0.476 | **0.879** |
| Decision Tree | C4.5 | **0.64** | **0.618** | **0.658** | **0.732** | **0.695** | **0.758** | **0.642** | **0.73** | **0.844** |
| | CART | **0.674** | **0.818** | **0.754** | **0.709** | **0.703** | **0.819** | **0.832** | **0.842** | **0.9** |
| | RF | **0.706** | **0.584** | **0.605** | 0.483 | 0.735 | **0.696** | **0.901** | **0.734** | **0.601** |

- Almost all classifiers that implemented GAFS+B method outperform the original method
- GAFS+B affected significantly on the performance of the class imbalance suffered classifiers

# Without GAFS+B vs With GAFS+B

| Classifiers | | *P* value of t-Test | Result |
|---|---|---|---|
| Statistical Classifier | LR | 0.156 | Not Sig. ($\alpha > 0.05$) |
| | LDA | **0.00004** | **Sig. ($\alpha < 0.05$)** |
| | NB | 0.294 | Not Sig. ($\alpha > 0.05$) |
| Nearest Neighbor | k-NN | **0.00002** | **Sig. ($\alpha < 0.05$)** |
| | K* | **0.001** | **Sig. ($\alpha < 0.05$)** |
| Neural Network | BP | **0.008** | **Sig. ($\alpha < 0.05$)** |
| Support Vector Machine | SVM | **0.03** | **Sig. ($\alpha < 0.05$)** |
| Decision Tree | C4.5 | **0.0002** | **Sig. ($\alpha < 0.05$)** |
| | CART | **0.0002** | **Sig. ($\alpha < 0.05$)** |
| | RF | **0.01** | **Sig. ($\alpha < 0.05$)** |

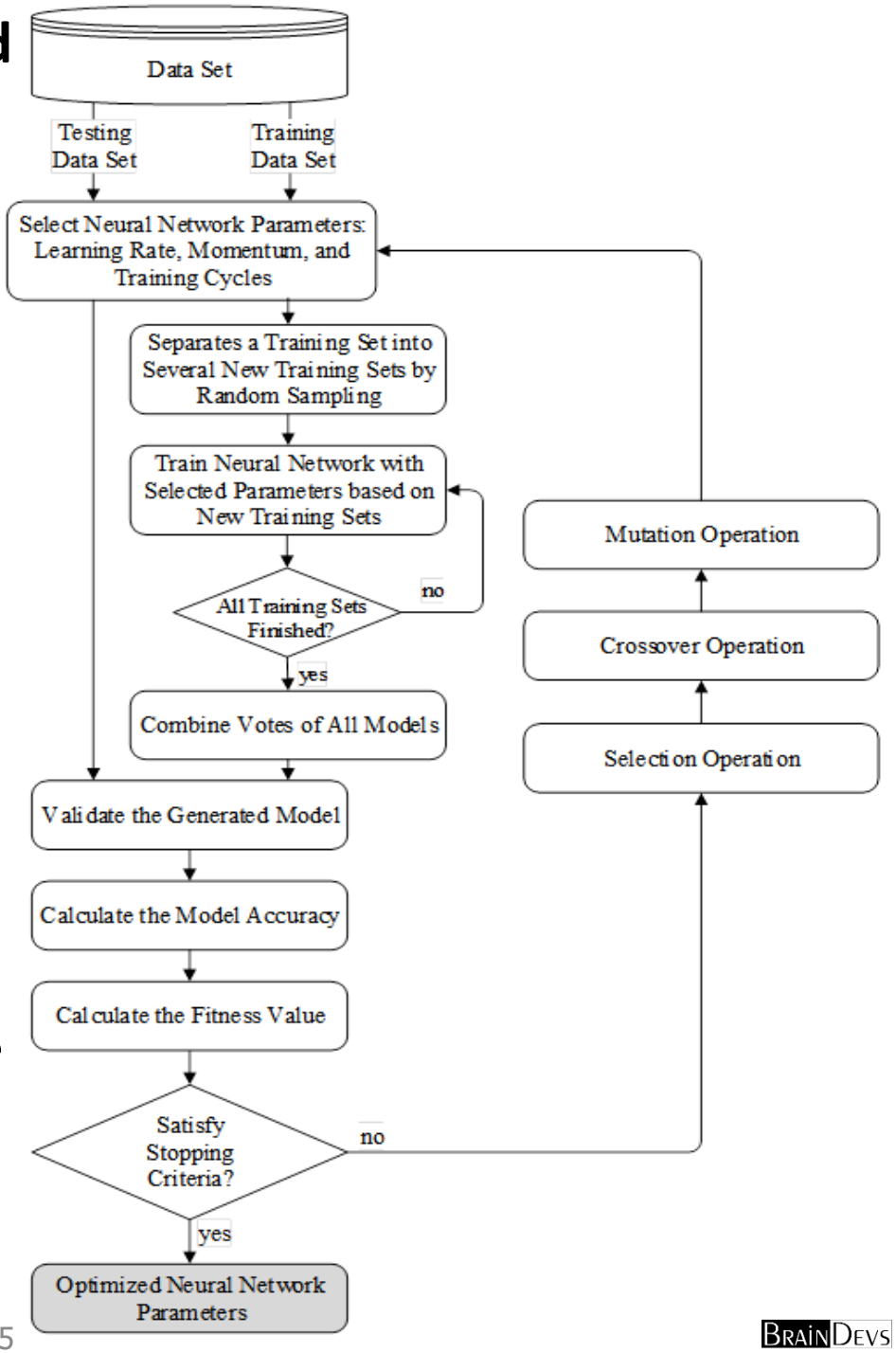
- Although there are two classifiers (LR and NB) that have no significant difference (*P value* > 0.05), the remaining eight classifiers (LDA, k-NN, K*, BP, SVM, C4.5, CART and RF) have significant difference (*P value* < 0.05)
- The proposed GAFS+B method makes an improvement in prediction performance for most classifiers

# A Hybrid Genetic Algorithm based Neural Network Parameter Optimization and Bagging Technique for Software Defect Prediction (NN GAPO+B)

- Every chromosome is evaluated by the fitness function Equation

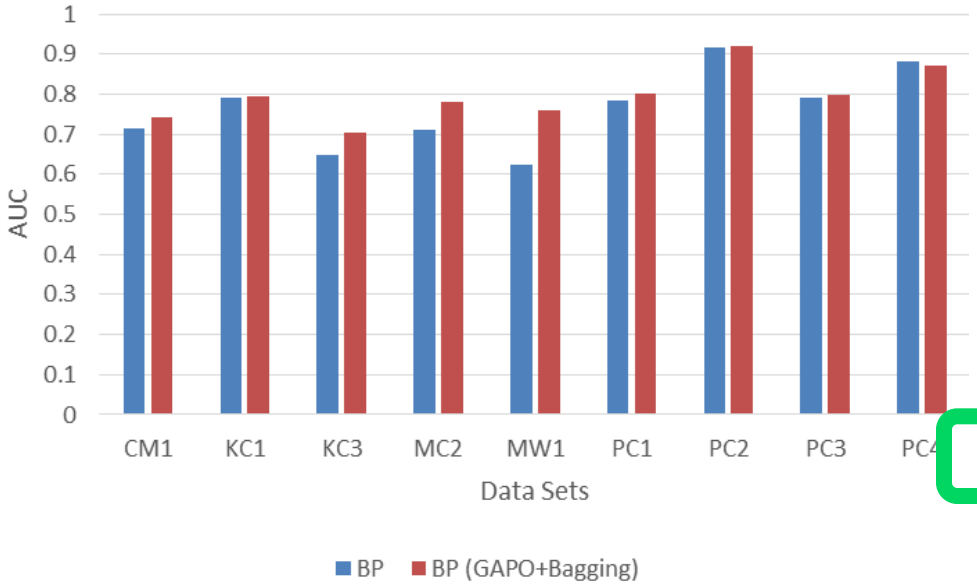
$$fitness = W_A \times A + W_P \times \left( S + \left( \sum_{i=1}^{n} C_i \times P_i \right) \right)^{-1}$$

- Where
  - $A$: classification accuracy
  - $P_i$: parameter value
  - $W_A$: weight of classification accuracy
  - $W_p$: parameter weight
  - $C_i$: feature cost
  - $S$: setting constant
- When ending condition is satisfied, the operation ends and the optimized NN parameters are produced. Otherwise, the process will continue with the next generation operation



# Results: NN GAPO+B

| Classifiers | CM1 | KC1 | KC3 | MC2 | MW1 | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|---|---|---|---|---|
| NN | 0.713 | 0.791 | 0.647 | 0.71 | 0.625 | 0.784 | 0.918 | 0.79 | 0.883 |
| **NN GAPO+B** | **0.744** | **0.794** | **0.703** | **0.779** | **0.76** | **0.801** | **0.92** | **0.798** | 0.871 |

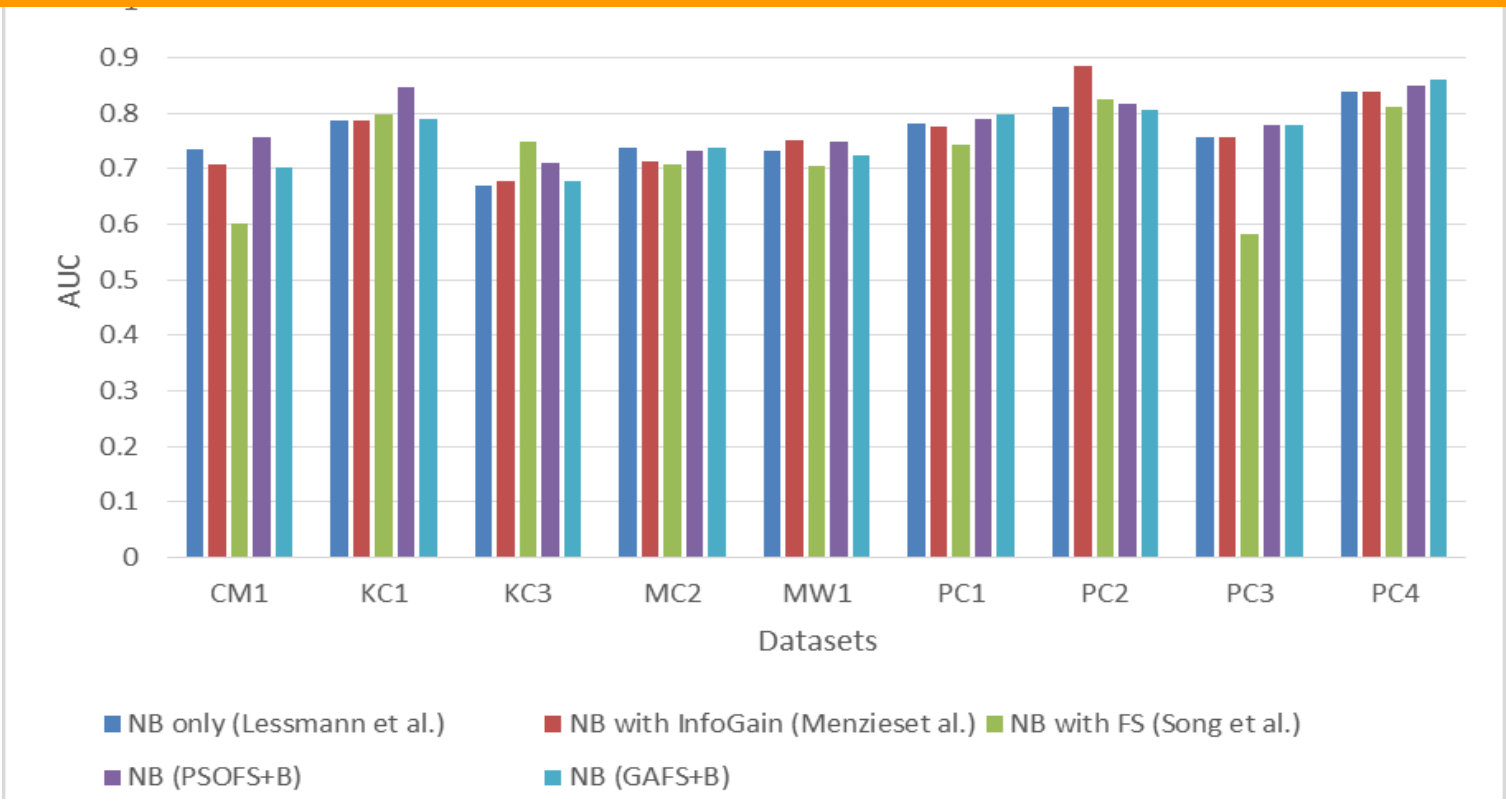



| | Variable 1 | Variable 2 |
|---|---|---|
| Mean | 0.762333333 | 0.796666667 |
| Variance | 0.009773 | 0.004246 |
| Observations | 9 | 9 |
| Pearson Correlation | 0.923351408 | |
| Hypothesized Mean Difference | 0 | |
| df | 8 | |
| t Stat | -2.235435933 | |
| ***P*** | **0.02791077** | |

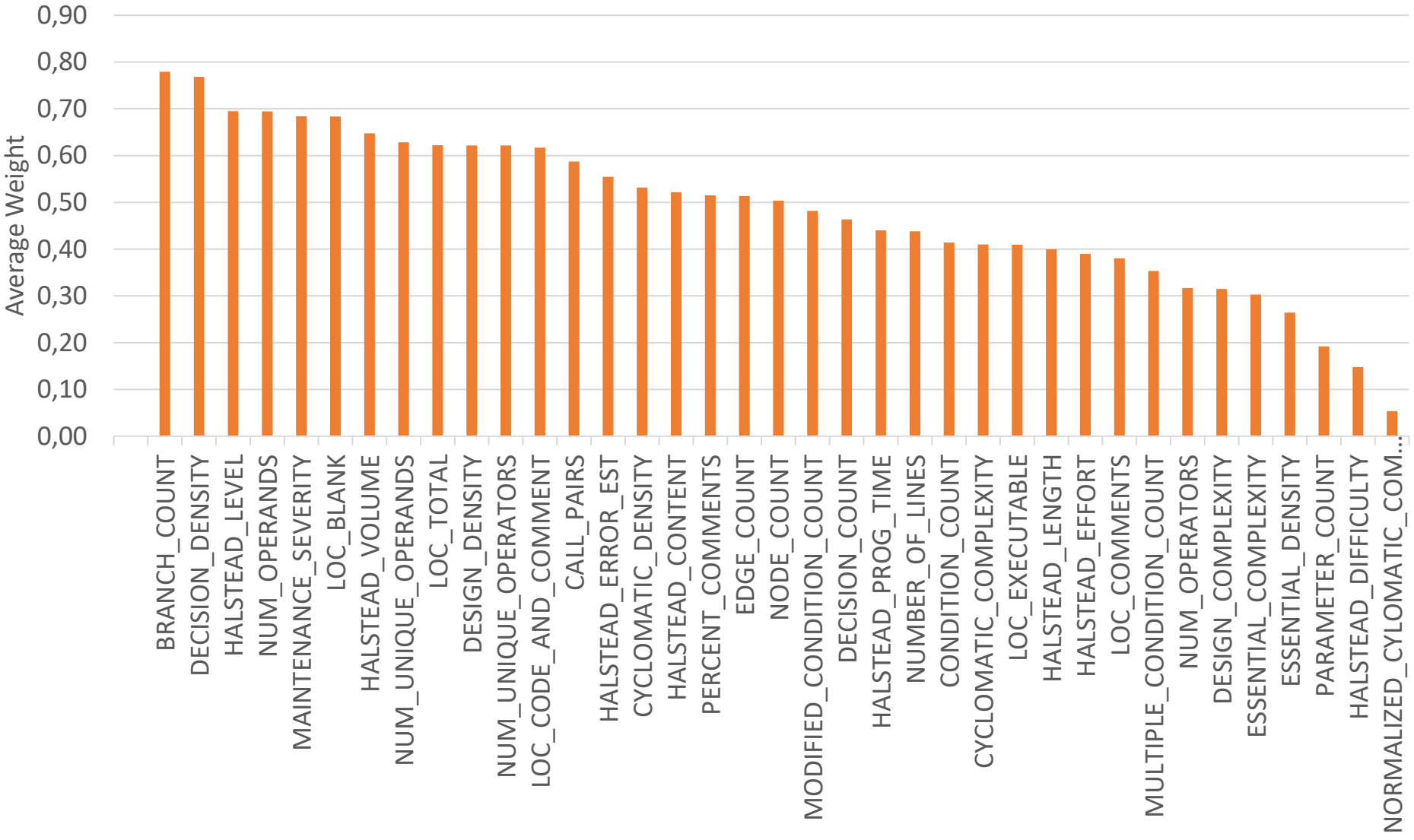
- NN GAPO+B outperforms the original method in almost all datasets
- The proposed (NN GAPO+B) method makes an improvement in prediction performance for back propagation neural network ($P<0.05$)

# Framework Comparison

| | CM1 | KC1 | KC3 | MC2 | MW1 | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|---|---|---|---|---|
| NB only (Lessmann et al.) | 0.734 | 0.786 | 0.67 | **0.739** | 0.732 | 0.781 | 0.811 | 0.756 | 0.838 |
| NB with InfoGain (Menzies et al.) | 0.708 | 0.786 | 0.677 | 0.712 | **0.752** | 0.775 | **0.885** | 0.756 | 0.84 |
| NB with FS (Song et al.) | 0.601 | 0.799 | **0.749** | 0.707 | 0.704 | 0.742 | 0.824 | 0.583 | 0.812 |
| NB (PSOFS+B) | **0.756** | **0.847** | 0.71 | 0.732 | 0.748 | 0.79 | 0.818 | **0.78** | 0.85 |
| NB (GAFS+B) | 0.702 | 0.79 | 0.677 | **0.739** | 0.724 | **0.799** | 0.805 | **0.78** | **0.861** |

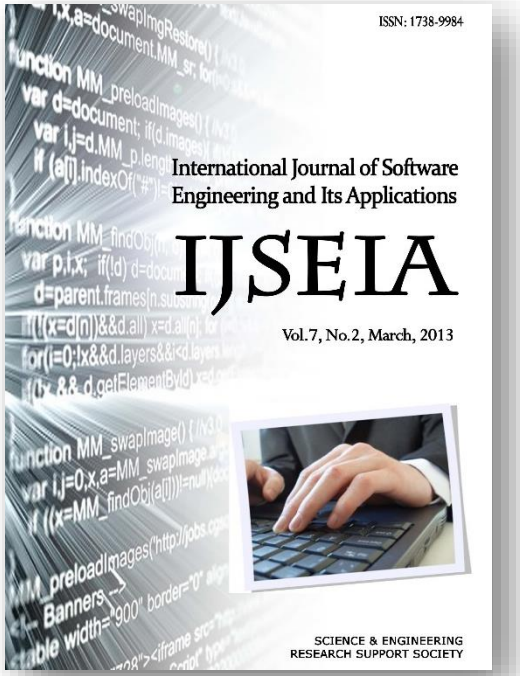


# Weighted Average of Relevant Attributes



# Research Publication on RQ3

Romi Satria Wahono and Nanna Suryana, *Combining Particle Swarm Optimization based Feature Selection and Bagging Technique for Software Defect Prediction*, International Journal of Software Engineering and Its Applications, Vol 7, No 5, September 2013



# Reference


- Abbott, M., & McKinney, J. (2013). Understanding and Applying Research Design. John Wiley & Sons, Inc.
- Berndtsson, M., Hansson, J., & Olsson, B. (2008). Thesis Projects: a Guide for Students in Computer Science and Information Systems (2nd ed.). London: Springer-Verlag
- Blaxter, L., Hughes, C., & Tight, M. (2006). How to Research (3rd ed.). Open University Press
- Blessing, L. T. M., & Chakrabarti, A. (2009). DRM, a Design Research Methodology. Springer-Verlag London
- Cohen, L., Manion, L., & Morrison, K. (2005). Research Methods in Education (5th ed.). Taylor & Francis Group.
- Dawson, C. W. (2009). Projects in Computing and Information Systems A Student's Guide (2nd ed.). Pearson Education Limited
- Jonker, J., & Pennink, B. (2010). The Essence of Research Methodology. Springer-Verlag Berlin Heidelberg
- Lichtfouse, E. (2013). Scientific Writing for Impact Factor Journals. Nova Science Publishers, Inc.

# Reference


- Kothari, C. (2004). Research Methodology: Methods and Techniques. New Age International
- Might, M. (2010). The Illustrated Guide to a Ph.D. Matt.might.net. Retrieved from *http://matt.might.net/articles/phd-school-in-pictures/*
- Marczyk, G., DeMatteo, D., & Fertinger, D. (2005). Essentials of Research Design and Methodology. John Wiley & Sons, Inc.
- Rea, L. M., & Parker, R. A. (2014). Designing and Conducting Survey Research: A Comprehensive Guide (4th ed.). John Wiley & Sons, Inc.
- Runeson, P., Host, M., Rainer, A., & Regnell, B. (2012). Case Study Research in Software Engineering: Guidelines and Examples. John Wiley & Sons, Inc.
- Sahu, P. K. (2013). Research Methodology: A Guide for Researchers In Agricultural Science, Social Science and Other Related Fields. Springer.
- Veit, R., Gould, C., & Gould, K. (2013). Writing, Reading, and Research (9th ed.). Cengage Learning.

# Terima Kasih

Romi Satria Wahono
*romi@romisatriawahono.net*
*http://romisatriawahono.net*
*08118228331*