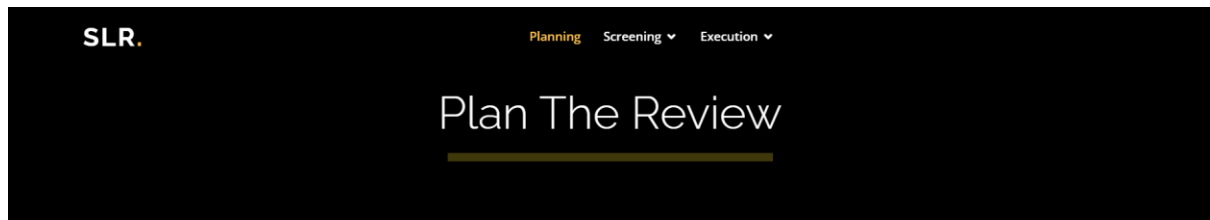


## How-To-Use Systematic Literature Review (SLR)

### Planning Page

1. Go to this link: <http://slr.aminhakim.tech>
2. Press on 'Download' button to download How-To Guide



### HOW-TO GUIDE

Please download the guide on how to use SLR:



### SELECTION CRITERIA

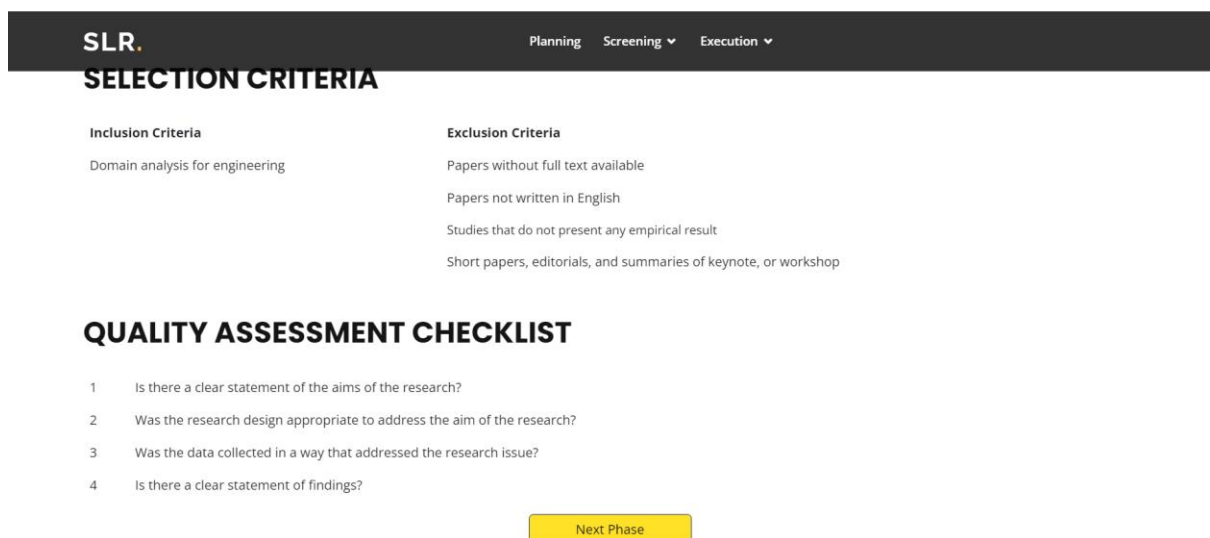
#### Inclusion Criteria

Domain analysis for engineering

#### Exclusion Criteria

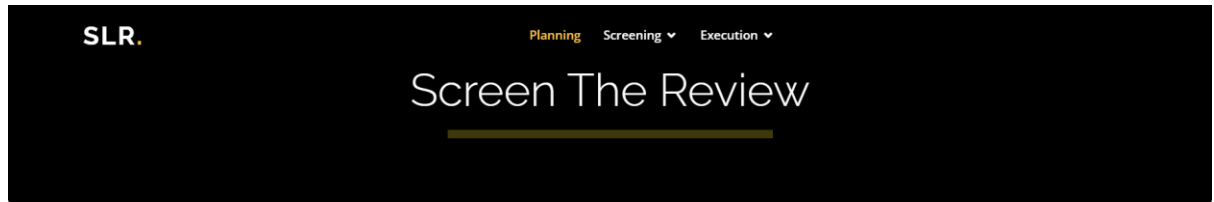
Papers without full text available

3. Click on 'Next Phase' button to proceed to the 'Screening' page



## **Screening Page**

1. User will be redirected to 'Screening' page from previous steps or can use this link:
2. Choose either Supervised Machine Learning or Unsupervised Clustering



**Choose your preferred approach for Filter Studies process:**

Supervised Machine Learning

Classify records as "relevant" or "not relevant" by training a machine learning model on a manually classified set of records

**Choose your preferred approach for Retrieve Studies process:**

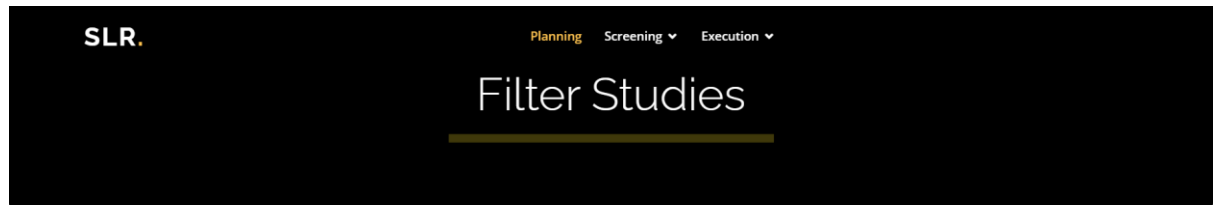
Unsupervised Clustering

Classify records as "relevant" or "not relevant" using a manually classified set of relevant records as seed values

IF choose 'Filter – Supervised Machine Learning

### Filter Page – Supervised Machine Clustering

1. User will be redirected to 'Filter' page from previous steps or can use this link:



### APPROACH: SUPERVISED MACHINE LEARNING

Machine Learning algorithms to assesses the probability of documents being relevant to a topic of interest

Machine Learning requires a training dataset (at least 25-100 documents; more training data will produce better results), annotated as relevant or not to the topic of interest.

#### IMPORT INPUT FILES

Upload Data (Standard CSV format) :

Choose file

Browse

2. Upload one file for 'Input File' and one file for 'Training File'
3. Click on 'Upload' button

**SLR. APPROACH: SUPERVISED MACHINE LEARNING**

PlanningScreeningExecution

Machine Learning algorithms to assesses the probability of documents being relevant to a topic of interest

Machine Learning requires a training dataset (at least 25-100 documents; more training data will produce better results), annotated as relevant or not to the topic of interest.

#### IMPORT INPUT FILES

Upload Data (Standard CSV format) :

Choose fileBrowse

#### IMPORT TRAINING FILES

Upload Data (Standard CSV format) :

Choose fileBrowse

Upload

### ALGORITHM SETTINGS

- Requirement for 'Input File'

	A	B	C	D	E
1	Title	Abstract	Year	Authors	Label
2	A Concept	This is an c	2006	Anicet Yala	1
3	A Quantita	<p>Requir	2007	Alan Davis	1
4	A survey a	<p>A comp	2007	Huzefa Ka	0
5	An analysis	OBJECTIVE	2005	Carolyn M	0
6	Challenges	Modeling i	2014	Michiel Re	1
7	Controver	This article	2007	M. N. Wicl	0
8	Data sets	OBJECTIVE	2005	Gernot A. I	1
9	Developing	Open sour	2015	Joseph Fel	1
10	Effectiven	This paper	2006	By Alan Da	1
11	Evidence-E	Several stu	2005	By Magne	1
12	Experimen	There is a	2005	Martin H?e	1

- Make sure the file format is .CSV
- Must has Abstract, Year, Authors, Label columns
- Make sure the titles of columns are correctly spelled
- Label column are filled as follows:
  - 1 – For relevant paper
  - 0 – For irrelevant paper

- Requirement for 'Training File'

	A	B	C	D
1	Title	Abstract	Year	Authors
2	A Concept	This is an c	2006	Anicet Yala
3	A Quantita	<p>Requir	2007	Alan Davis
4	A survey a	<p>A comp	2007	Huzefa Ka
5	An analysis	OBJECTIVE	2005	Carolyn M
6	Challenges	Modeling i	2014	Michiel Re
7	Controver	This article	2007	M. N. Wicl
8	Data sets	OBJECTIVE	2005	Gernot A. I
9	Developing	Open sour	2015	Joseph Fel
10	Effectiven	This paper	2006	By Alan Da
11	Evidence-E	Several stu	2005	By Magne
12	Experimen	There is a	2005	Martin H?e

- Make sure the file format is .CSV
- Must has Abstract, Year, Authors columns
- Make sure the titles of columns are correctly spelled

4. Select Supervised Machine Learning algorithm to use
5. Click on 'Run Machine Learning' button

SLR

Standard CSV format

PlanningScreeningExecution

Choose file

Browse

## IMPORT TRAINING FILES

Upload Data (Standard CSV format):

Choose file

Browse

Upload

## ALGORITHM SETTINGS

Topic extraction algorithms do not require any training data (i.e., no document annotation is needed), and are designed for rapid assessments to discover the major topics in a document collection.

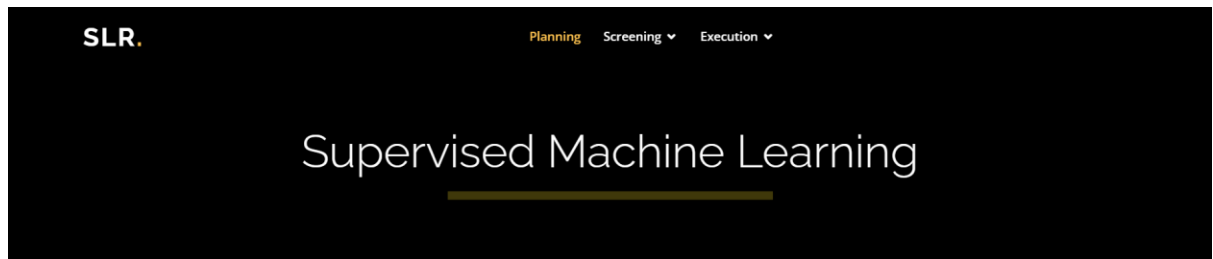
Select supervised technique:

Random Forest

Run Machine Learning

## **Filter Output Page**

1. User will be redirected to 'Filter Output' page
2. This page can only be accessed through 'Run Machine Learning' button from previous page
3. SLR will display the result of training the 'Training File' from previous page



## **RESULTS**

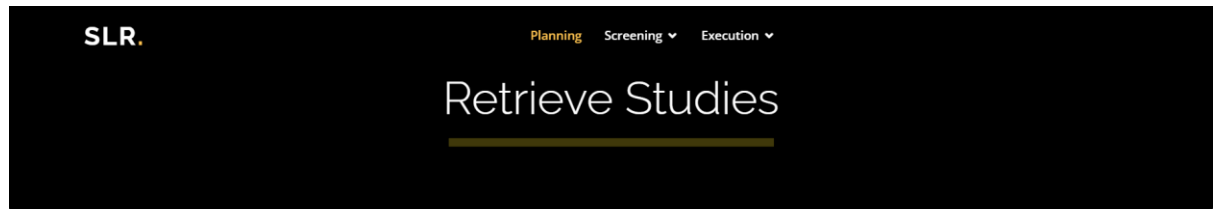
No.	Year	Title	Authors	Predict
0	2006	A Conceptual Model of ICT-Supported Unified Process of International Outsourcing of Software Production	Anicet Yalaho	1
1	2007	A Quantitative Assessment of Requirements Engineering Publications ? 1963? 2006	Alan Davis , Ann Hickey , Oscar Dieste , Natalia Juristo and Ana Moreno	1
2	2007	A survey and taxonomy of approaches for mining software repositories in the context of software evolution	Huzefa Kagdi, Michael L. Collard, Jonathan I. Maletic	1
3	2005	An analysis of data sets used to train and validate cost prediction systems	Carolyn Mair, Martin Shepperd, Magne J?rgensen	1

4. Click on 'Download' button to download the 'Training file' that has been trained with Machine Learning (as shown in the table)
5. This trained 'Training file' will be used for the next phase
6. Click on 'Next Phase' button to proceed to the 'Assess' page

IF choose 'Retrieve – Unsupervised Clustering'

## Retrieve Page – Unsupervised Clustering

1. User will be redirected to 'Retrieve page from previous steps or can use this link:



### IMPORT STUDIES

Upload Data (XLS format) :

Choose file Browse Upload

### ALGORITHM SETTINGS

Topic extraction algorithms do not require any training data (i.e., no document annotation is needed), and are designed for rapid assessments to discover the major topics in a document collection.

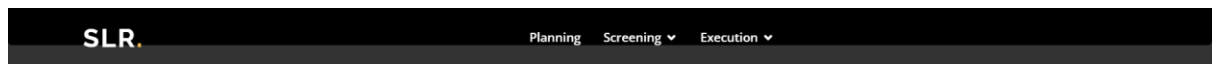
Select clustering technique:

2. Upload one file for 'Import Studies'
3. Click on 'Upload' button
  - Requirement for 'Import Studies' file

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Publication	Authors	Book Auth	Book Edit	Book Grou	Author Ful	Book Auth	Group Aut	Article Title	Source Tit	Book Serie	Book Serie	Language
2	C	Zaidi, MA		Gervasi, O		Zaidi, Moa			Conceptua	COMPUTA	Lecture Nc		English
3	J	Khakpour,				Khakpour,			Converger	TECHNOL			English
4	J	Kumeno, F				Kumeno, F			Software e	INTELLIGE			English
5	C	Sharma, P		Gupta, A; I		Sharma, P			Systematic	2017 INTE			English
6	J	Robledo, S				Robledo, S			Hasta la vi	JOURNAL			English
7	J	Nogales, F				Nogales, F			Hand gesti	INTERNAT			English
8	J	Kang, ZQ;				Kang, Ziqi			Machine le	COMPUTE			English
9	J	Henrique,				Henrique,			Literature	EXPERT S			English
10	J	van Klomp				van Klomp			Crop yield	COMPUTE			English
11	J	Saputri, TF				Saputri, TF			The Applic	IEEE ACC			English
12	J	Azeem, MI				Azeem, MI			Machine le	INFORMA			English
13	J	Korkmaz, I				Korkmaz, I			A review o	EDUCATIC			English
14	J	Bertolini, M				Bertolini, M			Machine L	EXPERT S			English
15	J	Alenezi, H				Alenezi, H			Utilizing cr	EDUCATIC			English
16	C	Mekkaoui,		Benadada		Mekkaoui,			A Systema	GOL'20: 20			English
17	J	Tsunoda, I				Tsunoda, I			Machine le	REVISTA			Portuguesi
18	C	Perez-Ver		JuarezRan		Manuel Pe			A Systema	2020 8TH			English
19	J	Wen, JF; I				Wen, Jian			Systematic	INFORMA			English

- Make sure the file is downloaded from Scopus and choose Export to Excel
- When exporting, choose 'Record Content' as Full Record
- Make sure the format is .XLS
- Do not change anything on the file downloaded
- Rename the file if needed

- SLR will display the result of cleaning files from paper that are missing either Year, Title or Abstract



## IMPORT STUDIES

Upload Data (XLS format) :

Choose file

Browse

Upload

The dataset uploaded is cleaned by removing records with missing fields and duplication. The summary of of process is shown for reference.

Description	Total Papers
Initial dataset	100
Remove missing year	8
Remove missing title	0
Remove missing abstract	0
The final dataset after removing of 8 papers	92

## ALGORITHM SETTINGS

- Select Unsupervised Clustering algorithm to use
- Click on 'Run Topic Extraction' button

SLR.		Total Papers	Planning	Screening	Execution
Initial Dataset		100			
Remove missing year		8			
Remove missing title		0			
Remove missing abstract		0			
The final dataset after removing of 8 papers		92			

## ALGORITHM SETTINGS

Topic extraction algorithms do not require any training data (i.e., no document annotation is needed), and are designed for rapid assessments to discover the major topics in a document collection.

Select clustering technique:

LDA

Number of Topics:

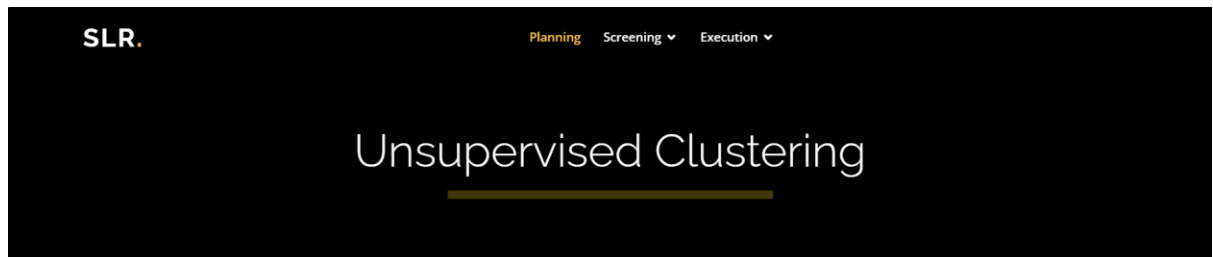
5

Run Topic Extraction



## **Retrieve Output Page**

1. User will be redirected to 'Retrieve Output' page
2. This page can only be accessed through 'Run Unsupervised Clustering' button from previous page
3. SLR will display the result of training the 'Import Studies' file from previous page



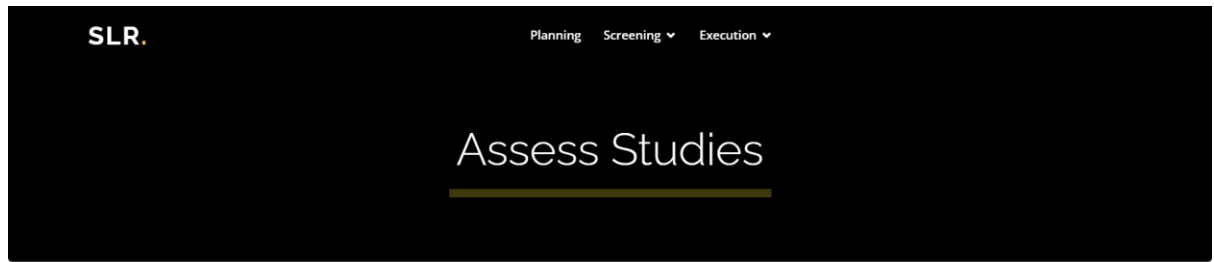
## **RESULTS**

No.	Year	Title	Authors	Topics
0	2021	Conceptual Modeling Interacts with Machine Learning - A Systematic Literature Review	Zaidi, MA	3
1	2021	Convergence of Gamification and Machine Learning: A Systematic Literature Review	Khakpour, A; Colomo-Palacios, R	2
2	2019	Software engineering challenges for machine learning applications: A literature review	Kumeno, F	0
3	2017	Systematic Literature Review on Software Effort Estimation Using Machine	Sharma, P; Singh, J	3

4. Click on 'Download' button to download the 'Import Studies' file that has been trained with Unsupervised Clustering (as shown in the table)
5. This trained 'Import Studies' file will be used for the next phase
6. Click on 'Next Phase' button to proceed to the 'Assess' page

## Assess Page

1. User will be redirected to 'Assess' page from previous steps or can use this link:



2. Upload one file for 'Quality Assessment'
3. Click on 'Upload' button

IF choose 'Filter' – Supervised Machine Learning

- Make sure the format is .CSV

IF choose 'Retrieve – Unsupervised Clustering

- Make sure the format is .XLSX

### Requirement for 'Quality Assessment' file

	A	B	C	D	E	F	G
1	Index	Title	Abstract	Year	Authors	Predict	Relevant
2	0	A Concept	This is an c	2006	Anicet Yala	1	1
3	1	A Quantita	<p>Requir	2007	Alan Davis	1	1
4	2	A survey a	<p>A comp	2007	Huzefa Kag	1	1
5	3	An analysis	OBJECTIVE	2005	Carolyn M	1	0
6	4	Challenges	Modeling i	2014	Michiel Re	1	1
7	5	Controver	This article	2007	M. N. Wicl	1	1
8	6	Data sets	OBJECTIVE	2005	Gernot A.	1	0
9	7	Developing	Open sour	2015	Joseph Fel	1	1
10	8	Effectiven	This paper	2006	By Alan Da	1	0
11	9	Evidence-E	Several stu	2005	By Magne	1	0
12	10	Experimen	There is a	2005	Martin H?s	1	0

- DO NOT delete any column
- Add 'Relevant' column and fill the column as follows
  - 1 – For Relevant paper
  - 0 – For Irrelevant paper
- Make sure the 'Relevant' column is correctly spelled

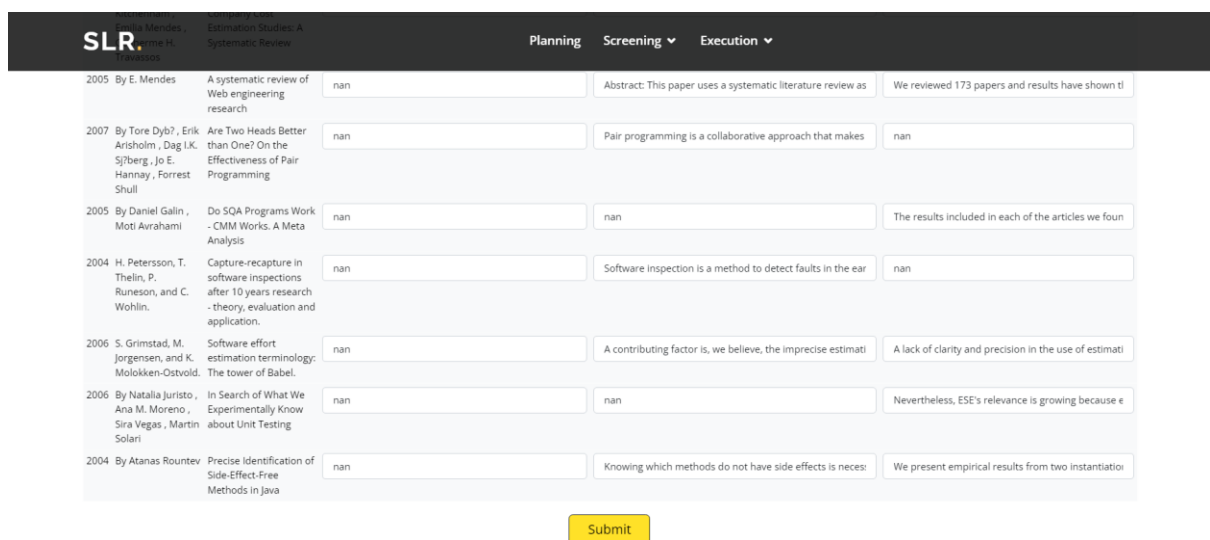
## Extract Data Page

1. User will be redirected to this page from the previous page.
2. Only paper that are labelled as 'Relevant' or 1 will be shown
3. Choose Objective, Method, and Result for each paper.



Year	Author	Title	Objective	Method	Result
2006	Anicet Yalaho	A Conceptual Model of ICT-Supported Unified Process of International Outsourcing of Software Production	This study proposes the information communication t	At the theoretical level, the model can be used as a basis f	To address this research question, the results of ar
2007	Alan Davis , Ann Hickey , Oscar Dieste , Natalia Juristo and Ana Moreno	A Quantitative Assessment of Requirements Engineering Publications ? 1963? 2006	It is important to recognize the plethora of results acc	nan	It is important to recognize the plethora of results.
2007	Huzefa Kagdi, Michael L. Collard, Jonathan I. Maletic	A survey and taxonomy of approaches for mining software repositories in the context of software evolution	A taxonomy is derived from the analysis of this literati	<p>A comprehensive literature survey on approaches for	nan

4. Click on 'Submit' button to proceed to the next phase.



## Synthesis Page

1. User will be redirected to this page from the previous page
2. SLR will show the Word Cloud for Objective, Method, and Result



## RESULT

### Word Cloud - Objective



3. SLR will show the Pie and Bar Chart for selected papers
4. User will be able to use 'Download' button to download all the diagrams displayed.

