



How to Choose a Project?

Data Science and AI/ML

Dr. Manoj Jayabalan

Agenda



1

Introduction

2

Systematic Process

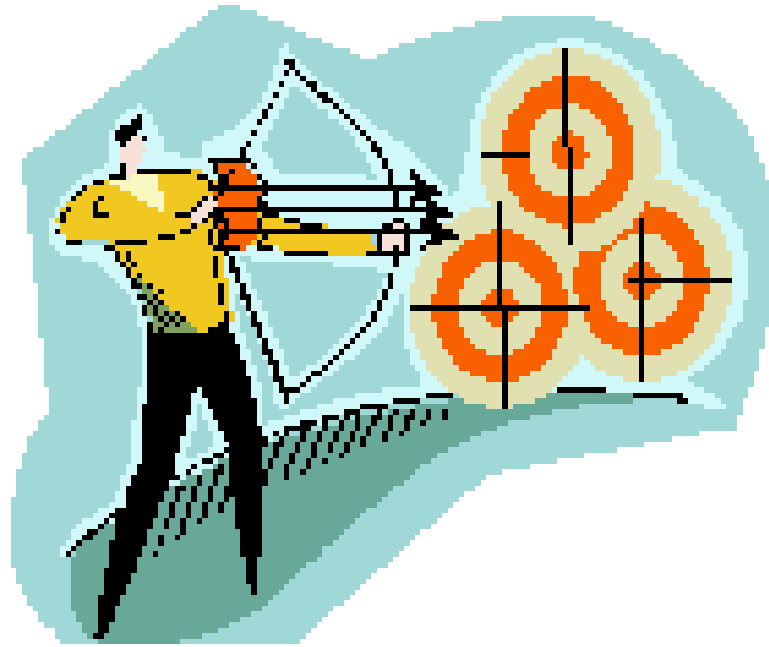
3

Exercise

4

Q & A

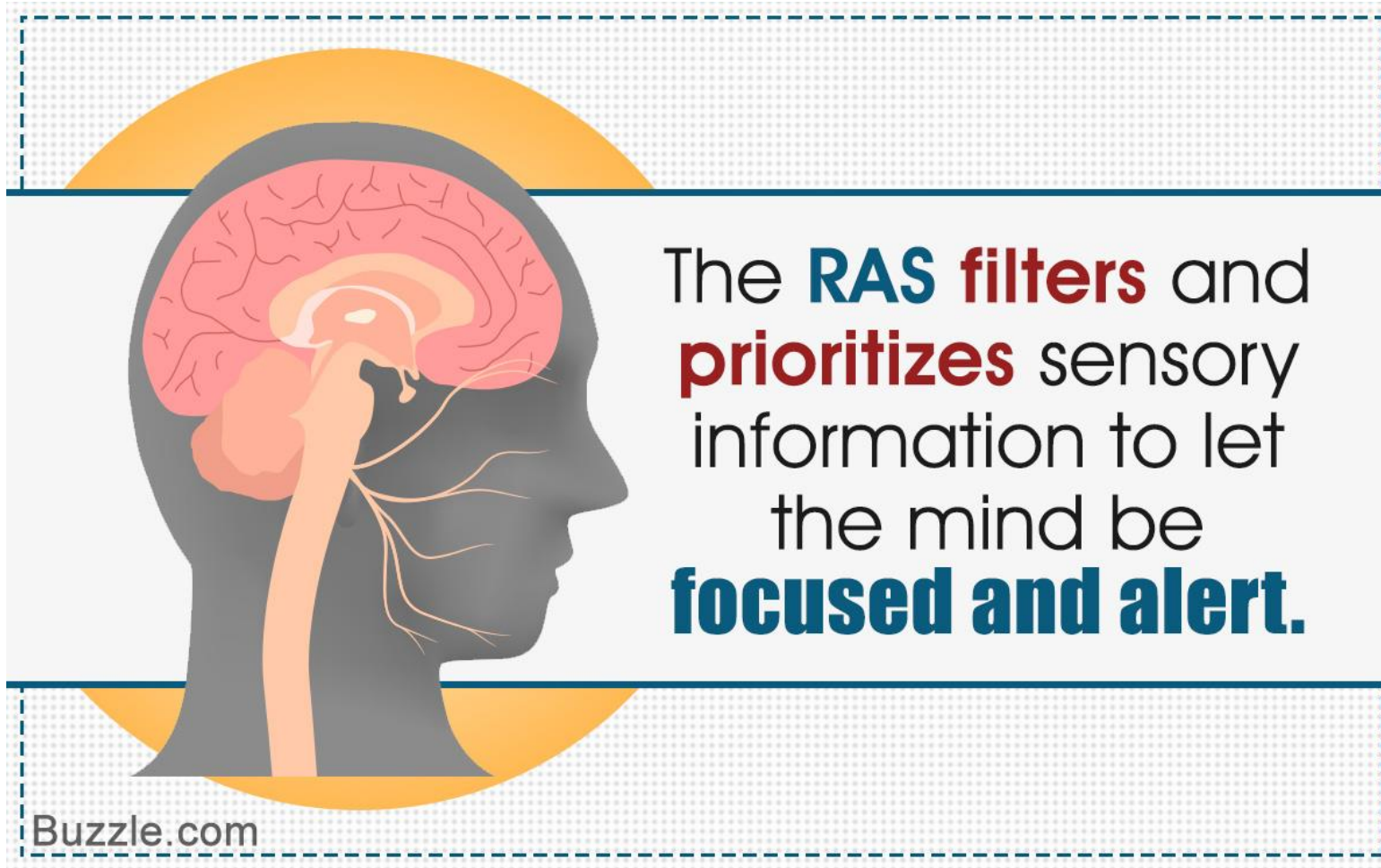
Project Title



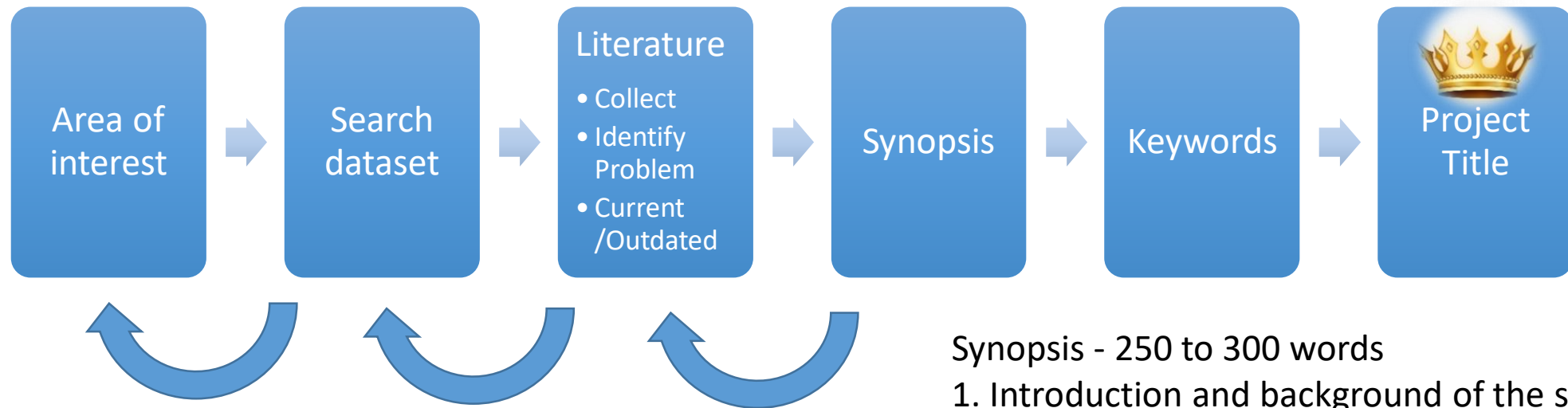
© Copyright: technicalpathways.com

*A clear project title means
one target, one arrow.*

Have you heard of Reticular Activating System (RAS) in your brain ?



Systematic Process



Synopsis - 250 to 300 words

1. Introduction and background of the study
2. Problems
3. Methods
4. Purpose
5. Expected results
6. Conclusions

Exercise

1. Area of Interest

Education Data Mining

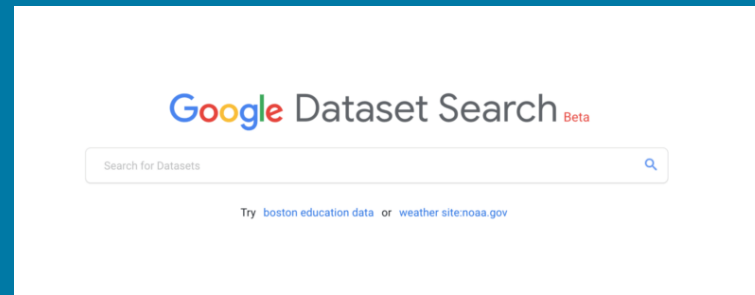
- Program recommendation
- Student performance
- Student dropout
- Student retention ...etc



Exercise

2. Search Datasets

Education Data Mining – “Student Performance”



kaggle



IEEE DataPort™



Exercise – Additional dataset source

2. Search Datasets



**Academic
Torrents**



Google Cloud



data.world



Research Center

Exercise

3. Literature

Education Data Mining – “Student Performance”



Machine Learning Repository

Center for Machine Learning and Intelligent Systems

Student Performance Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: Predict student performance in secondary education (high school).

Data Set Characteristics:	Multivariate	Number of Instances:	649	Area:	Social
Attribute Characteristics:	Integer	Number of Attributes:	33	Date Donated	2014-11-27
Associated Tasks:	Classification, Regression	Missing Values?	N/A	Number of Web Hits:	529435

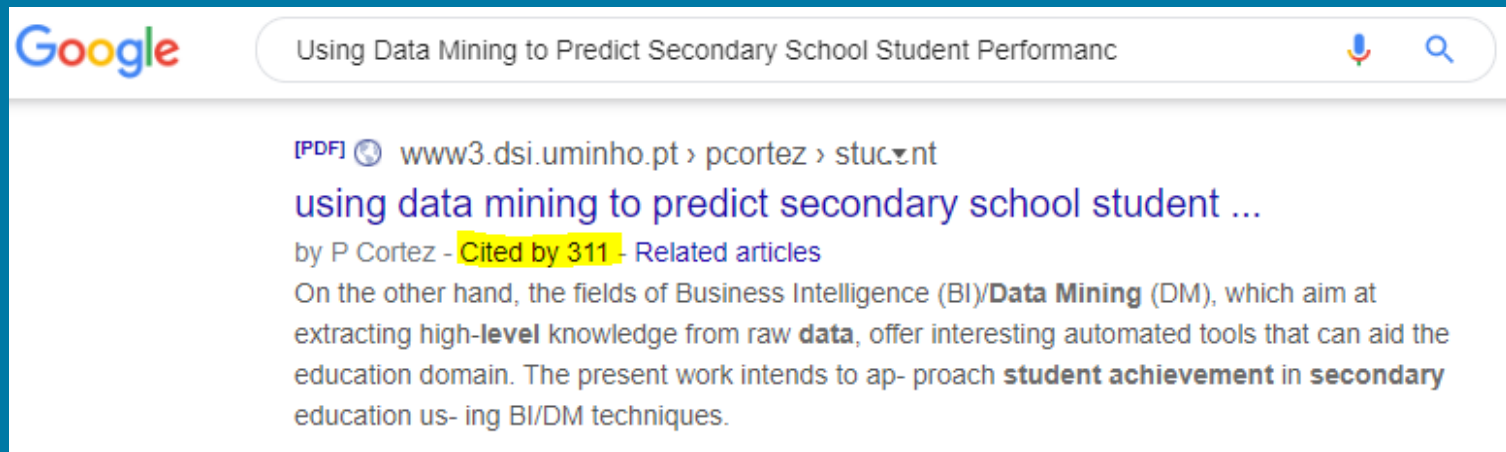
Source:

Paulo Cortez, University of Minho, Guimarães, Portugal, <http://www3.dsi.uminho.pt/pcortez>

Exercise

3. Literature

Education Data Mining – “Student Performance”



The screenshot shows a Google search interface. The search bar contains the text "Using Data Mining to Predict Secondary School Student Performanc". Below the search bar, the first result is a PDF document titled "using data mining to predict secondary school student ..." by P Cortez. The document is cited by 311 and has related articles. The abstract of the document is visible, discussing the application of Business Intelligence (BI) and Data Mining (DM) techniques to predict student achievement in secondary education.

Google

Using Data Mining to Predict Secondary School Student Performanc

[PDF] www3.dsi.uminho.pt › pcortez › student

using data mining to predict secondary school student ...

by P Cortez - Cited by 311 - Related articles

On the other hand, the fields of Business Intelligence (BI)/Data Mining (DM), which aim at extracting high-level knowledge from raw data, offer interesting automated tools that can aid the education domain. The present work intends to approach student achievement in secondary education using BI/DM techniques.

Exercise

3. Literature

IEEE
Xplore[®]
Digital Library

ScienceDirect

arXiv.org

 Springer

 Google
Scholar

ACM  DIGITAL
LIBRARY

PubMed

 ResearchGate

Exercise

4. Synopsis



Adobe Acrobat
Document

Education Data Mining – “Student Performance”

Read the given document,
extract the keywords and suggest a
suitable title.

Exercise

**5. Keywords
&
6. Title**

Education Data Mining – “Student Performance”

What is the title?

Q & A

“Every project is an opportunity to learn, to figure out problems and challenges, to invent and reinvent.”

- David Rockwell

Read more at <https://www.brainyquote.com/topics/project-quotes>