

Intelligent Reasoning Systems Project

- **Project Title:** Pre-school Recommender
- **Group Number (Registered in Canvas):** XXX
- **Group Members:**
 - Hu Lei (A0120681N)
 - Ho Zi Hao Timothy (A0150123B)
 - Tan Jinn Yeu (A0067140N)
 - Pan Lifan (A0291192A)

Introduction

This project is about building a **recommender system** to aid parents in searching for a suitable pre-school **based on their preferences and constraints**.

The application can help parents in **simplifying** the decision and **demystifying** what to look out for in a pre-school in today's world of exploding information on the internet such as forums, pre-school websites, social media. Currently, parents are limited to internet research, site visits or basic search tool for preschools.

The goal of the project is to build a **Minimum Viable Product (MVP)** to test out whether the model recommendation based on limited set of features is useful to a parent with child of pre-school going age before considering to expand its features and scope.

Project Background

The Singapore government is increasingly concerned about ensuring that every child does not **fall behind** in early childhood education, given its importance.¹ However, parents are faced with the **tradeoffs** of wanting the best for their child and the reality of dual income family and high cost of living in Singapore, despite the government support.

Underlying this trend is that young parents today are more digitally savvy and are also relying on digital media for information and social interactions, as compared to the older generations. With the explosion of information on the internet, it is becoming **harder to distill out relevant information** to make the best decision to which pre-school best meets their needs and situation. This is where the recommender system can **support** in their decision making.

¹ [Commentary: How Budget 2024 sets out to tackle wage gaps and social mobility in Singapore - TODAY \(todayonline.com\)](https://www.todayonline.com/singapore/commentary/how-budget-2024-sets-out-to-tackle-wage-gaps-and-social-mobility-in-singapore)

Project Background

Landscape

- **Problem faced by parents**

- Multiple suitable preschool factors to consider – distance, cost, center's values, teaching approaches, safe and quality center programmes, hygiene ², teachers, meals, enrollment and financial matters⁵
- Presence of different accrediting metrics – license tenure, length of tenure previously attained, SPARK accreditation ⁶ (not compulsory)
- Require domain knowledge to make optimal choice – price may not be best indicator of suitability (head, heart and gut factors³, expert advice)

- **Market Landscape**

- Large offering of preschool choices - 1,600 childcare centres and close to 380 kindergartens¹
- Manual research process by parents to search internet sites or make recce visits
- Few technology sites offering comparison, recommending or matching
- Lower level of technology innovation in preschool education information space

^{1,2,3} How do parents pick the right preschool? Use head, heart and gut, say experts - <https://www.todayonline.com/singapore/how-pick-right-preschool-experts-2250981>

⁴ 7 things to consider when choosing a preschool, from location to licence tenure - <https://www.channelnewsasia.com/singapore/choosing-preschool-what-look-out-tips-advice-3749441>

⁵ Choosing a preschool for your child - [ECDA | Choosing a preschool for your child](https://ecda.gov.sg/Choosing-a-preschool-for-your-child)

⁶ [spark-certified-centres_2jan2024.pdf](https://ecda.gov.sg/spark-certified-centres_2jan2024.pdf) (ecda.gov.sg)

Product-Market Fit Pyramid

Target User:	Underserved Needs:	Value Proposition:	Feature Set:
<ul style="list-style-type: none">• Parents of child who can receive early childhood education• Digitally savvy	<ul style="list-style-type: none">• Need to have a complete view of all pre-schools' offerings• Need to know what to consider in a pre-school• Unaware of how to decide which preschool is better suited for child and parent preferences	<ul style="list-style-type: none">• A solution to demystify and simplify pre-school selection	<ul style="list-style-type: none">• User input for personal preferences and constraints• Informed search and ranking of pre-schools meeting criteria from user input• Display of pre-school ranking

Porter's Five Forces Analysis of Product

Threat of New Entrants (Barriers to Entry): Medium

As the information is publicly available, it can be easy for a new entrant to build a similar application and roll it to market. However, network effect is hard to replicate once the application gains traction.

Threat of Substitutes: High

Buyers can choose not to use the application and rely on their own effort to collate information even though it might be less optimal.

Bargaining Power of Suppliers: Low

The IT infrastructure will be on cloud, making it easy to switch from one provider to another.

Bargaining Power of Buyers: Medium

There are alternate suppliers in the market today (e.g., LifeSG, Skoolpedia) even though there are lesser features.

Industry Rivalry: Low

There is no alternate supplier in the market today.

Competitor Analysis

	LifeSG	Skoolopedia	Preschooler
Brief Description	Government app for existing preschools	Website search on preschool	Web search for preschool
Search Tool	Yes	Yes	Yes
User reviews	No	Yes	Yes
Listing of features	Yes	Yes	Yes
Expert Review	No	Yes	No
Recommendation	No	No	No
Sentiment Analysis	No	No	No

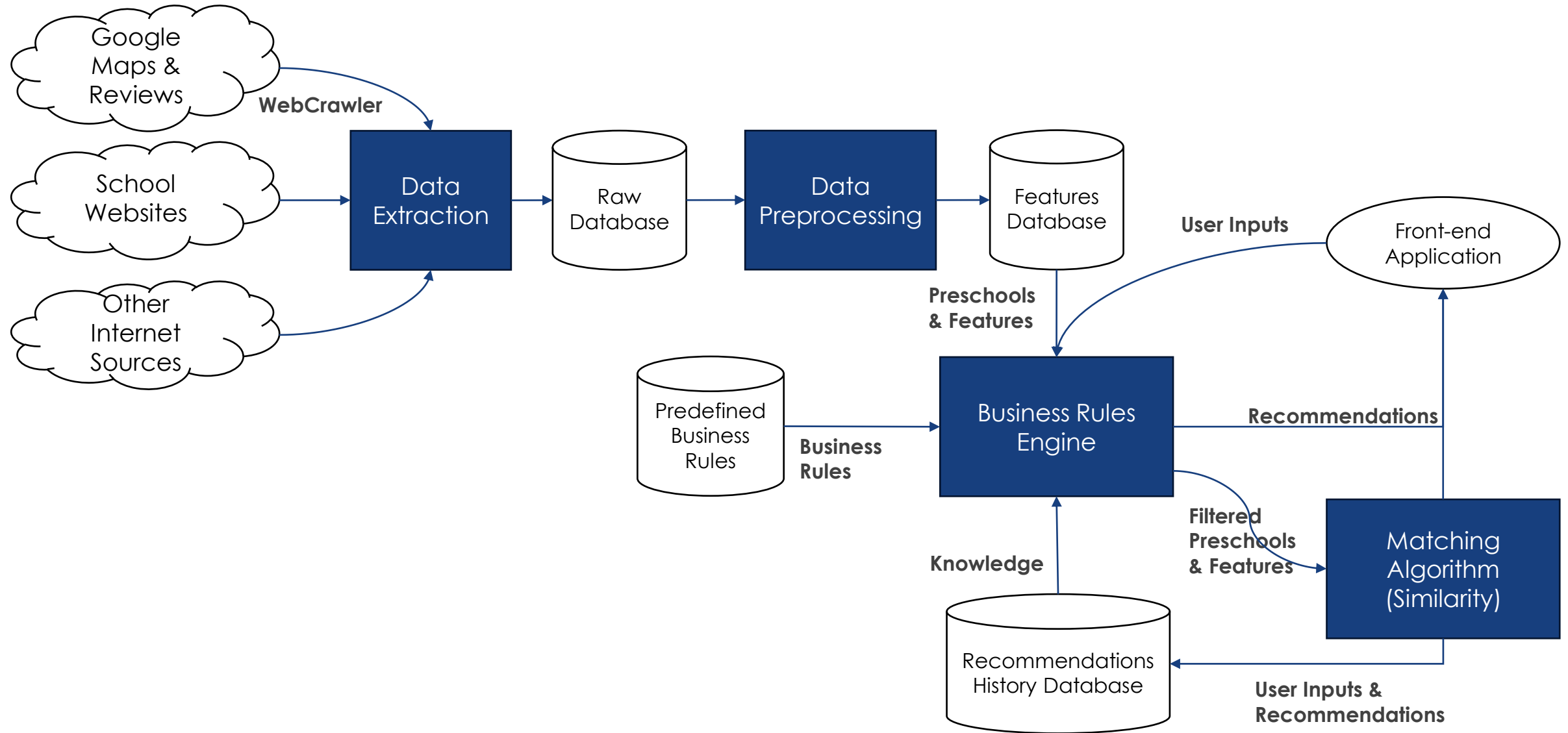
SWOT Analysis of Product

Strength <ul style="list-style-type: none">• This application relies on more features to derive a recommendation, compared to the competitors.• Summarises information for busy parents so they can make the right decision for selecting school.• Apply machine learning techniques to optimise decision making for parents with conflicting priorities in selecting schools.	Weakness <ul style="list-style-type: none">• This application is not a must-have for parents.• Parents may not be aware of such recommender apps to assist them.• Parents may prefer visiting preschool personally or rely on personal judgements to understand if learning environment is suitable.
Opportunities <ul style="list-style-type: none">• This application can be a sales lead generation tool for pre-school operators.• This application can be a tech-for-public-good product.	Threats <ul style="list-style-type: none">• This application can be replaced by a mega-app that integrate other aspects of parenting.• Some preschools may have poor online presence or outdated websites.

Project Scope

- 1) A **Web Scrapping Application** to capture & store information from the internet (e.g., Reviews).
- 2) A **Data Preprocessing Algorithm** to perform data cleaning and data transformation (e.g., one-hot encoding, normalization, sentiment analysis)
- 3) A **Front-end Application** for users to:
 - Provide input on (i) Preferences (e.g., Location) and its relative importance, (ii) Constraints (e.g., Budget).
 - View pre-school recommendation.
- 4) A **Business Rules Engine** to:
 - Decide if additional machine learning is required.
 - Filter out the features database based on business rules before passing it on for machine learning.
- 5) A **Matching Algorithm** based to rank the preschools based on similarity to the user inputs

System Design



Project Scope

Intelligent Reasoning Techniques

Requirement Group	Intelligent Reasoning Techniques
Decision automation: Business rules & process OR Knowledge based reasoning techniques	<p>Rule-based reasoning</p> <ul style="list-style-type: none">Filtering based on constraints of attending preschool (e.g., budget, distance) <p>Case-based reasoning</p> <ul style="list-style-type: none">Retrieval of past recommendations based on matching current user inputs and previous user inputs
Knowledge discovery & (big) data mining techniques	<p>Analogical machine learning</p> <ul style="list-style-type: none">Clustering & finding similarity between user inputs and preschool features
System designed with cognitive techniques or tools	<p>Knowledge base</p> <ul style="list-style-type: none">Storing of previous recommendations as knowledge to be used for knowledge-based reasoning

- 1) There is a trade-off between having a comprehensive criteria list and not cluttering the user interface. Therefore, a balance must be struck to filter out the key criteria. Consumer research should be periodically conducted to ensure that the criteria maintained are relevant and make a material impact on the decisions made.
- 2) Preschools may open and close without updating the website promptly. The data will need to be timestamped based on last refresh to ensure it is updated. A feedback channel should also be built to continually ensure the data quality.

Data Collection and Preparation

Data Sources

For facts on pre-schools, data sources include:

- 1) Pre-school Official Websites (e.g., MapleBear, MindChamps, My First Skool, etc.)
- 2) Preschool Certifications (e.g, [spark-certified-centres_2jan2024.pdf](#) ([ecda.gov.sg](#)))
- 3) Competitors' website (e.g., [Skoolopedia](#) , [preschooler](#))

For reviews on pre-schools, data sources include:

- 1) Parenting Forums Websites (e.g., [KiasuParents](#), Google Reviews)

For information on criteria looked out by parents, data sources include:

- 1) Traditional Media (e.g., commentary sections)
- 2) Information learnt through interviews with parents

Data Collection and Preparation

Challenges to Acquisition

To acquire the data sources mentioned, a web scrapping application is to be built to scrap from the websites.

No.	Possible Challenges	Mitigating Measures
1	As these data sources are websites, the websites data may be protected by anti-scraping mechanism (e.g., CAPTCHA blockers) to deter bots from crawling the sites.	To prevent triggering the anti-scraping mechanism, I. Read and understand the website's terms of service and robots. txt file. II. Limit the frequency of your scraping.
2	Some websites are single page applications and does not reveal the html pages.	Usage of ChromeDriver and Python Selenium / Playwright to overcome single page applications.

Data Collection and Preparation

Challenges to Processing

To process the data extracted from the websites, a preprocessing algorithm is to be built to find the relevant information based on key search terms (e.g., price, \$)

No.	Possible Challenges	Mitigating Measures
1	The terms used by each website may be different.	Perform natural language processing on the unstructured data (e.g., lemmatization) to find the root word.
2	Each website would also have a different website structure changes, putting different information on different pages.	Rely on site maps to classify the data extracted from the webpages. Start the search based on similar pages.

THANK YOU