COM4101 SENIOR PROJECT

INTELLIGENT MOBILE APP FOR DAILY USE

Market.AI

Group 11
Lau Ka Pui s226064
Mak Wai Hang s207900
Shum Wing Lam s207902
Tong Yuen Yu s207905

MARKET.AI

Download and install our app!

TABLE OF CONTENTS

01 MARKET ANALYSIS

MARKET SIZE OF STOCK TRADING AND INVESTING APPLICATIONS : USD 24.1 BILLION

Futubull-Easy Investment

Release in 2013

Investing.com: Stock Market

Release in 2013

Yahoo Finance – Stock Market

Release in 2015

02 OUR APP

THE MARKET.AI APP

DEVELOPMENT IDEA

- Demands of free chart pattern tools increased base on Global Market needs
- Providing a clearer information via Customizable financial tool that assess data of historical trends & stock prices

GOALS

- Provide users more information to study the stock historical trends
- Assist users to analyze the stock market

LICENSE/REGULATIONS

The following activities are exempted according to The Regulations of Hong Kong & SFC :

- Providing generic factual market information
- Providing analytical tools

- DO NOT need a HKMA/SFC license
- NO investment opinions & predictions

COMPARISON

	AI Analysis	Prediction	Trend Matching + Subsequent trends Analysis	Financial News
	yes	yes	no	yes
	no	no	no	yes
	no	no	no	yes
	yes	no	yes	yes

03 FEATURES

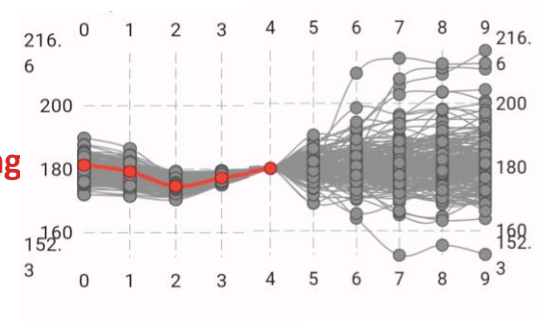
MAIN FEATURES

01

Candlestick charts

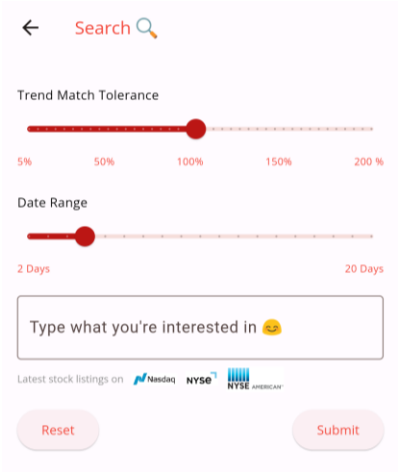
02

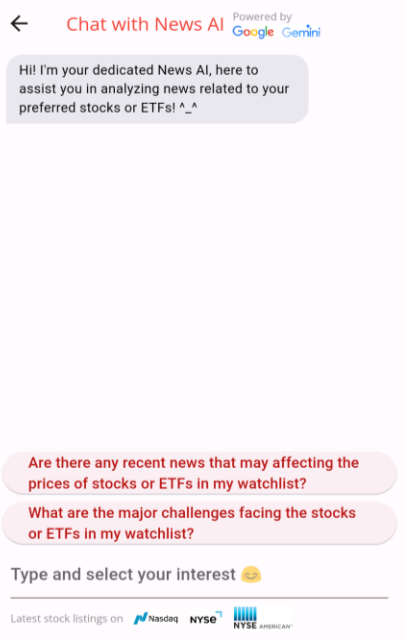
Historical trend matching



03

Subsequent trend analysis





MAIN FEATURES

04

Customized search

05

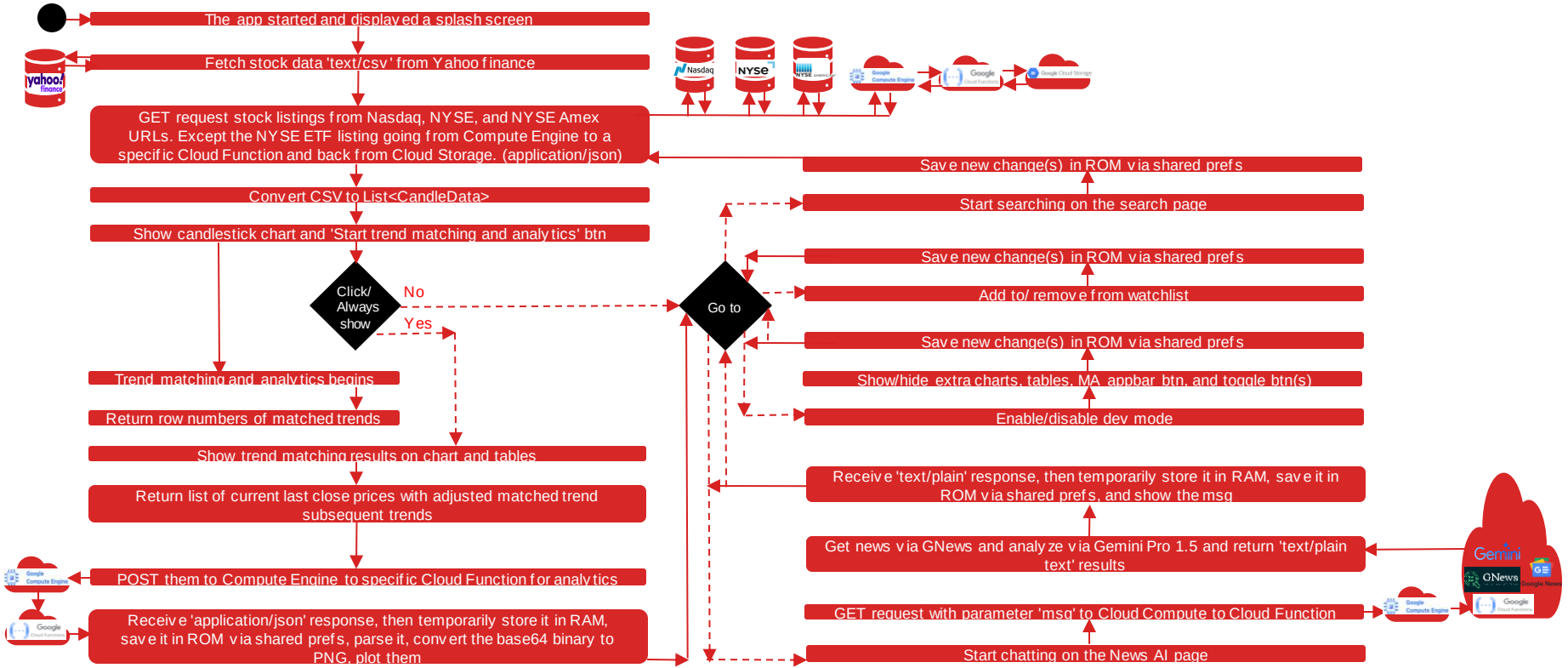
Generative AI aided news provider

04

DEVELOPMENT DETAILS

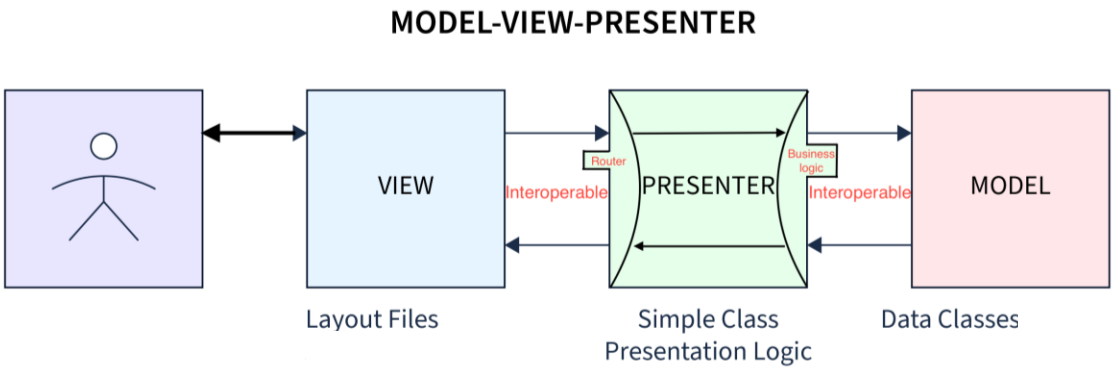
Architecture & concpets

ACTIVITY DIAGRAM



SOFTWARE DESIGN PATTERNS USED

- MVP Architecture

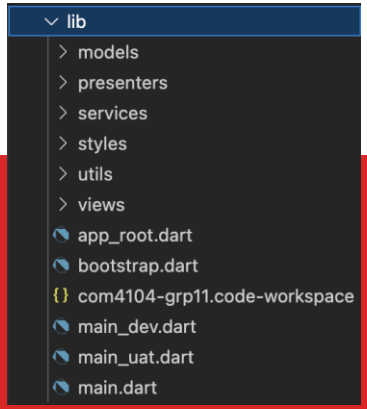


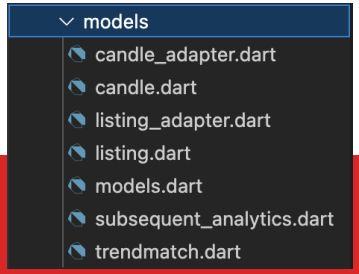
views
- charts
- chat.dart
- herophotoview_routewrapper.dart
- main.dart
- search.dart
- subsequent_analytics_view.dart
- trendmatch_view.dart
- views.dart

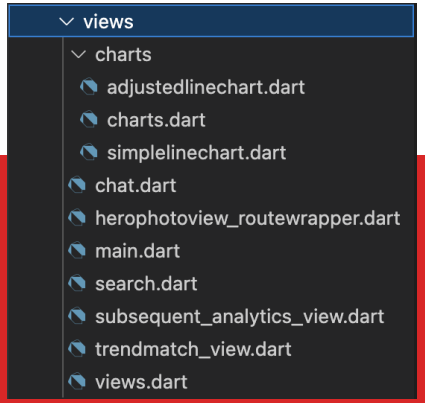
presenters
- main.dart
- presenters.dart

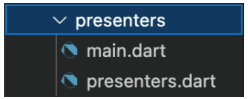
models
- candle_adapter.dart
- candle.dart
- listing_adapter.dart
- listing.dart
- models.dart
- subsequent_analytics.dart
- trendmatch.dart

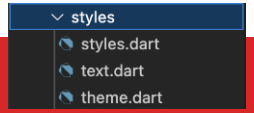
PROJECT STRUCTURE

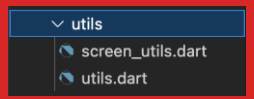


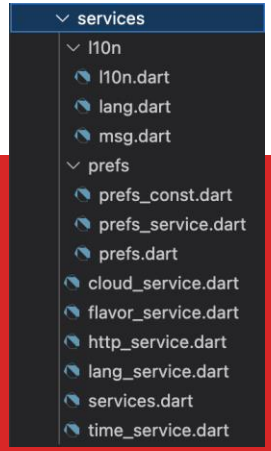












SOFTWARE HARDWARE ARCHITECTURAL

Minimum requirement:

OS	Android	iOS
Version	7.0	12
Hardware	Screen size: 5-inch 1080p 441ppi	
Network	Dial-up at 500 Kbps, possible with some performance loss (4G)	

Recommended requirement:

OS	Android	iOS
Version	12.0	16.1
Hardware	Screen size: 6.69-inch 2796×1290 460 ppi	
Network	Dial-up at 1 Mbps, possible with some performance loss (5G)	

CANDLESTICK CHART

- **Flutter interactive_chart package**
- **Data from Yahoo Finance**

https://query1.finance.yahoo.com/v7/finance/download/$stockSymbol?period1=$startTimestamp&period2=$endTimestamp&interval=1d&events=history&includeAdjustedClose=true

Check if current UTC time is greater than or equal to 20:00 in Eastern Daylight Time (USA summer and spring seasons)

Check if current UTC time is greater than or equal to 21:00 in Eastern Standard Time

→CSV →List<List<dynamic>> →List<CandleData> by adapter

Support:
- Pan
- Pinch-to-zoom
- Overlay
- Hide the overlay line (in developer mode)

yahoo! finance

		DL(ms)	Exe(ms)
AAPL.csv	733 KB	638	45
spy.csv	578 KB		

LISTINGS FROM EXCHANGES

- Data from Nasdaq, NYSE, NYSE Amex

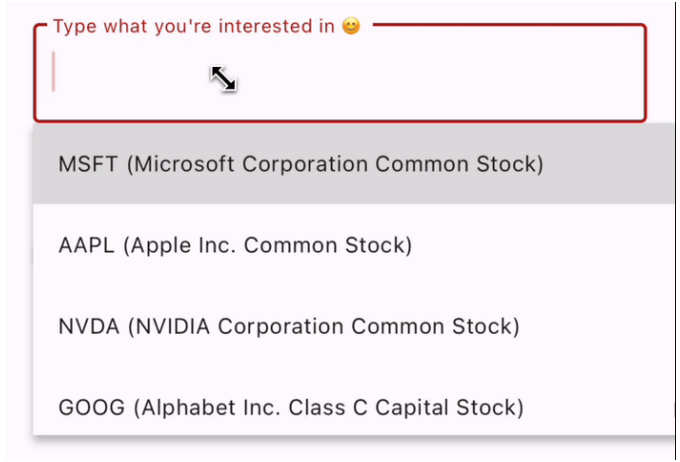
Stocks:
https://api.nasdaq.com/api/screener/stocks?tableonly=true&limit=99999&exchange=NASDAQ

https://api.nasdaq.com/api/screener/stocks?tableonly=true&limit=99999&exchange=NYSE

https://api.nasdaq.com/api/screener/stocks?tableonly=true&limit=99999&exchange=AMEX

ETFs:
http://35.221.170.30/?func=nyse-etfs-symbol-name

→text/json →List<SymbolAndName> by adapter

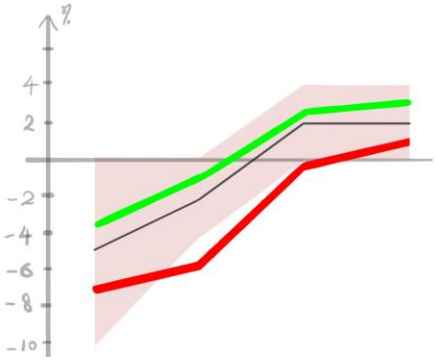


We have these listings to support our search utility

TREND MATCHING (METHOD)

Identification of matched trend:

- Selected last 5 days stocks
- Calculate the daily difference between the close price
- The close-price-differences compose a time series like [-5%, -2%, +2%, +2%]
- Look for all similar trends in the history of this stock
- Pseudo code:
    - For selected date
      Calculate the percentage change of closed price between the date and the next date (selList)
    - For every date in the history except selected date
      Calculate the percentage change of closed price between the date and the next date (comList)
    - For comList
      Calculate the percentage difference between selList and comList for 5 times(selected days) (pd)
      If all pd < 100% (Tolerance)
      Match trend + 1
      else
      Not Match trend + 1

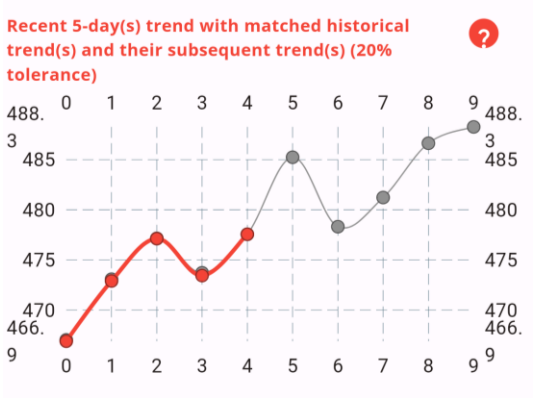


Match

Not match

TREND MATCHING (RESULT)

Challenge:
How similar (by %) = matched?

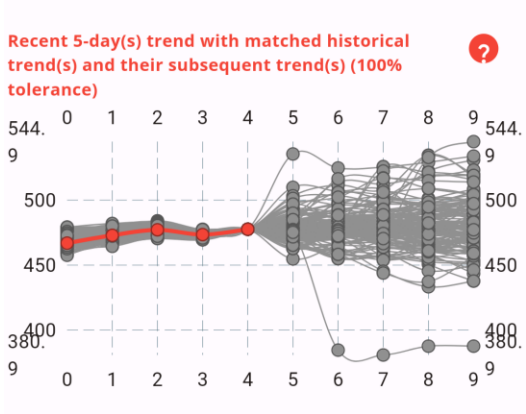


- Tolerance: the deviation from the daily difference time series
  - e.g., tolerance = 20%

    [-5%, -2%, +2%, +2%] is matched with
    [-6% ~ -4%, -2.4% ~ -1.6%, +1.6% ~ +2.4%, +1.6% ~ +2.4%]

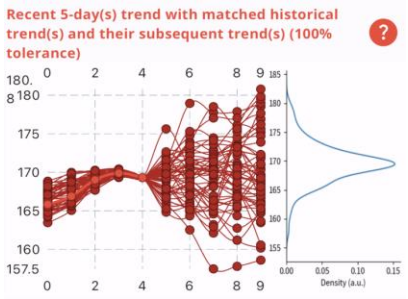
- Problem: inadequate match cases due to small tolerance
- Solution: enlarge the tolerance to 100%
  - e.g., tolerance = 100%

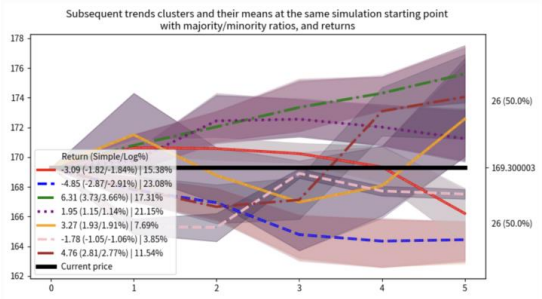
    [-5%, -2%, +2%, +2%] is matched with
    [-10% ~ 0%, -4% ~ 0%, 0% ~ +4%, 0% ~ +4%]

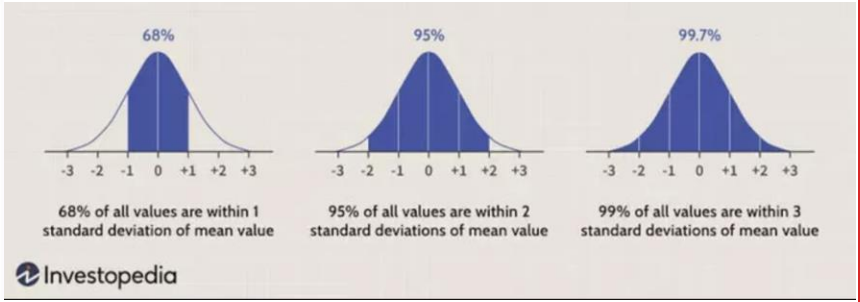


SUBSEQUENT TREND ANALYSIS - OBJECTIVE

- Three types of analysis:
  - Expected return by mean/median (by close price only) – center of the last day
  - Distribution with mean and +/- std (by close price only) – spread / shape of the last day (Applied 'The Empirical Rule')
  - Trend clustering (knowing main trends) – hidden main pattern of the time series
    - More informative
    - User can easily understand the main trends in complex trend distributions and why they are so similar, and they can easily determine which cluster fits the Subsequent trend.







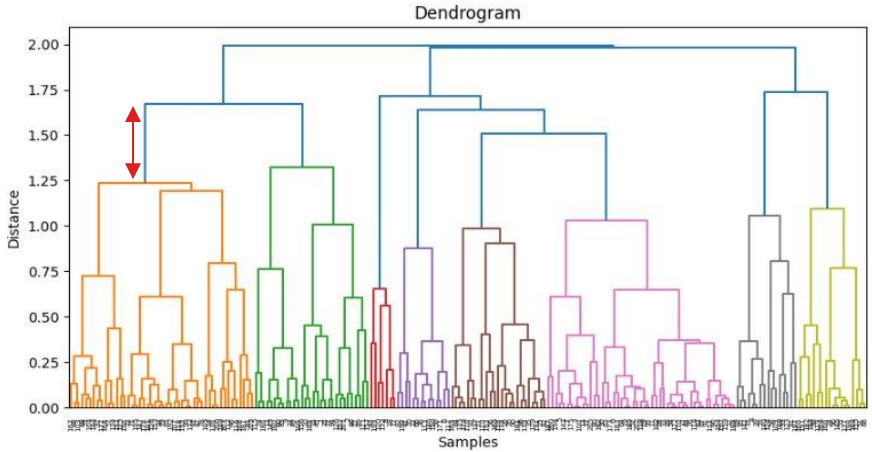
SUBSEQUENT TREND ANALYSIS – TIME SERIES CLUSTERING

Clustering algorithm implementation

- HAC (Hierarchical Clustering)
- Complete Linkage Clustering
- Correlation based

$$D(X,Y) = \max_{x \in X, y \in Y} d(x,y)$$

- x and y are the most distant elements in X and Y respectively
- Bottom-up approach
  - Each data represent one cluster → a single cluster



How many pairs should we use after calculation ?

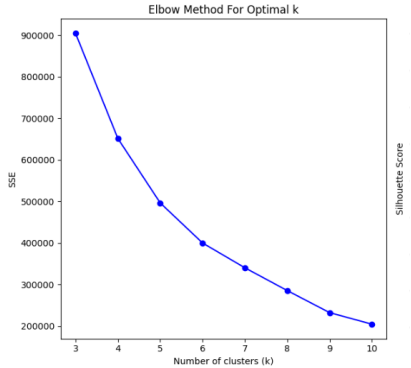
- **(next page -- solution)**

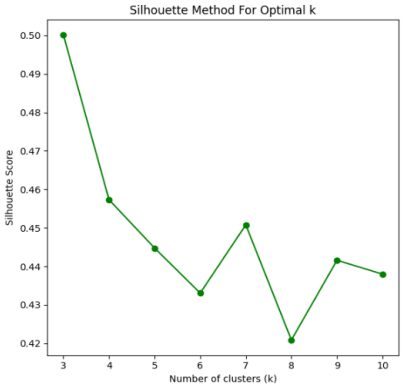
SUBSEQUENT TREND ANALYSIS – HOW MANY CLUSTERS?

- How many clusters?
- Elbow
  - Sometimes the elbow is not clear (intra distance only)
  - Elbow position is subjective
- Silhouette Coefficient

$$s = \frac{b-a}{\max(a,b)}$$

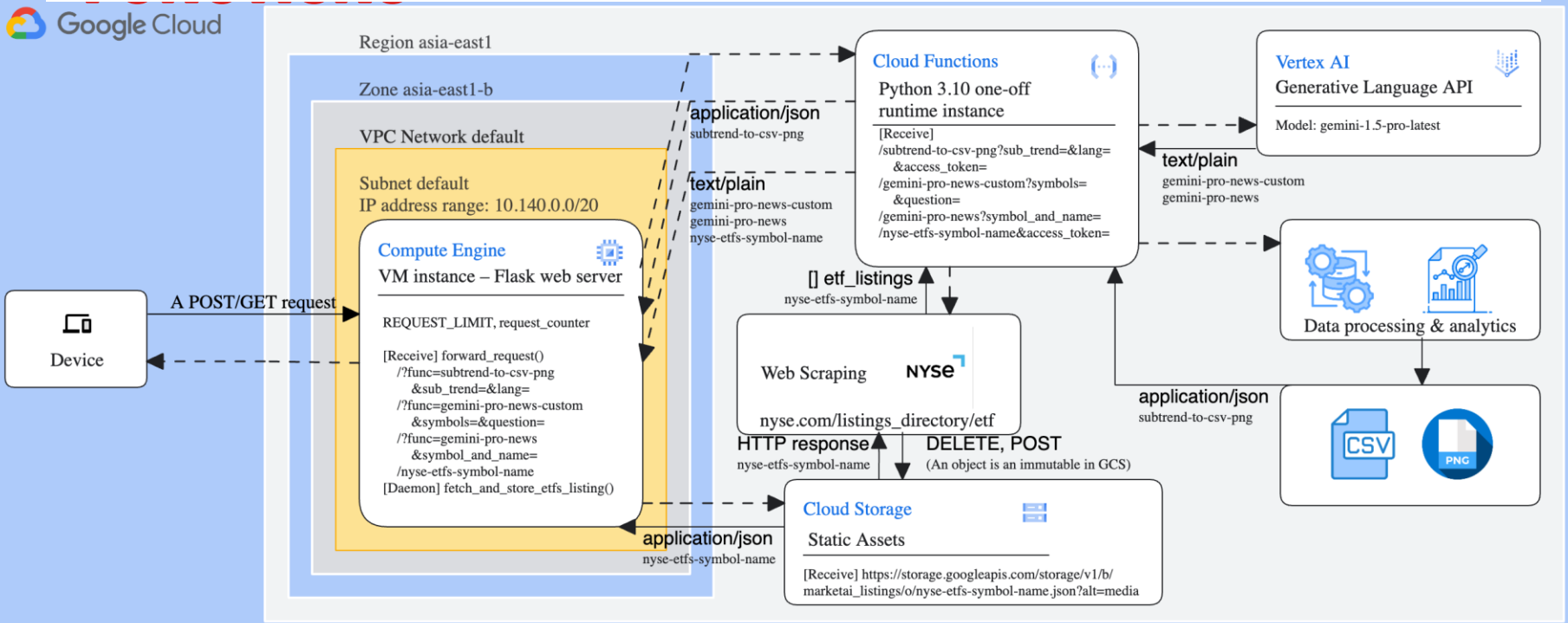
  - b is inter cluster distance
  - a is intra cluster distance
  - Consider both inter and intra cluster distance
  - There is always a max value
  - The weighting between b and a is subjective (we use 1 and 1 here)





SUBSEQUENT TREND ANALYSIS – CLOUD

- Clustering involves intricate computational processes that are not feasible to execute on the client side of an application. This is particularly true for Flutter, as it lacks packages for data mining and visualization.



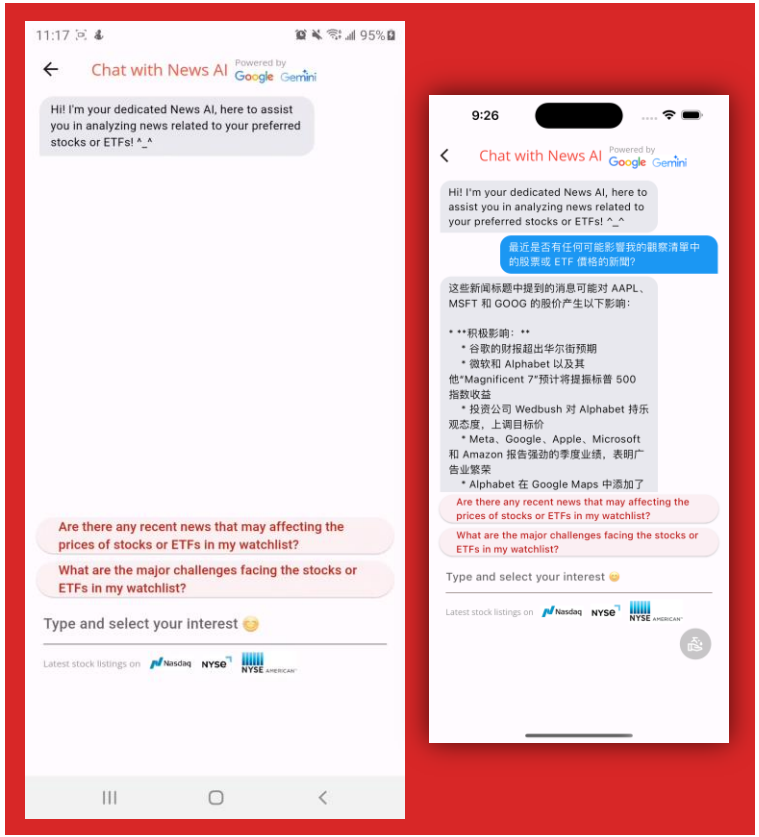
GENERATIVE AI AIDED NEWS PROVIDER

Reporting related news of financial instrument:

- Using GNews package, summarize a kinds of news from the entire internet, but we fine-tune it for our financial usage by Gemini Pro 1.5

In _ChatViewState class

- _sendMessage method -> **sending message to the Cloud Function**
- adding the message to the list of messages, updates the scroll position, and retrieves news analytics based on the message using the CloudService class
- **Cloud Function make API calls [Generative API]** & retrieve data
- aiResponseTime -> storing calculated response time of AI service



GENERATIVE AI AIDED NEWS PROVIDER

/gemini-pro-news-custom

```
# Questions:
# 1. Are there any recent news that may affecting the prices of
# 2. What are the major challenges facing the
res = MODEL.generate_content(question +
                             symbols + '?' +
                             ' Your answers should based on the following news titles only: ' +
                             str(news),
                             generation_config=LONG_GENERATION_CONFIG,
                             stream=False,
                             )
```

/gemini-pro-news

```
res = MODEL.generate_content('Are there any recent market news that may affecting the price of '
                             + symbol_and_name +
                             '? Your answers should based on the following news only:' +
                             str(news),
                             generation_config=LONG_GENERATION_CONFIG,
                             stream=False,
                             )
```

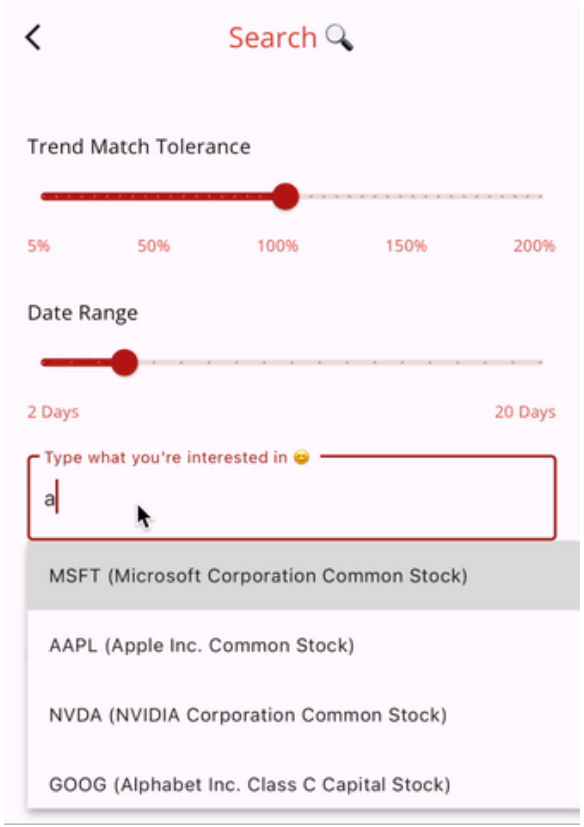
```
LONG_GENERATION_CONFIG = genai.GenerationConfig(
    temperature=0.9,
    top_p=1.0,
    top_k=32,
    candidate_count=1,
    max_output_tokens=8192,
)
```

- **temperature (0.1)**: Controls randomness in output. Lower = more deterministic.
- **top_p (1.0)**: Sets threshold for cumulative probability for word selection.
- **top_k (32)**: Limits next word selection to top K probable words.
- **candidate_count (1)**: Number of sequences generated before final output.
- **max_output_tokens (32)**: Maximum length of generated sequence.

CUSTOMIZED SEARCH

Allowing users to input the preferred date range, tolerance and stock
Generating graphs & data for users

- _displayStringForOption for formatting the display string for **autocomplete options (Entering stocks name & symbol)**
- _resetForm for **resetting the search form**
- _submitForm for **submitting the search form and performing validations**
- using the Get package and shared preferences (PrefsService) for **storing and retrieving user preferences temporarily or permanently**
- updating the UI



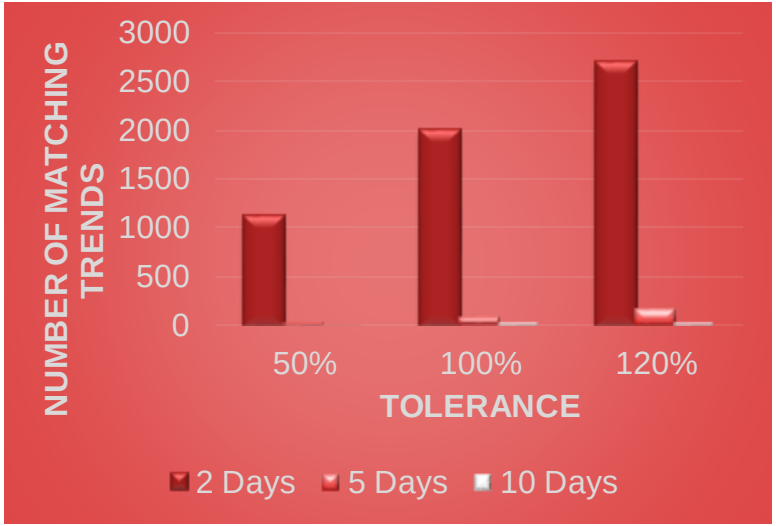
05

ANALYSIS & TEST CASES

SUBSEQUENT TREND ANALYSIS

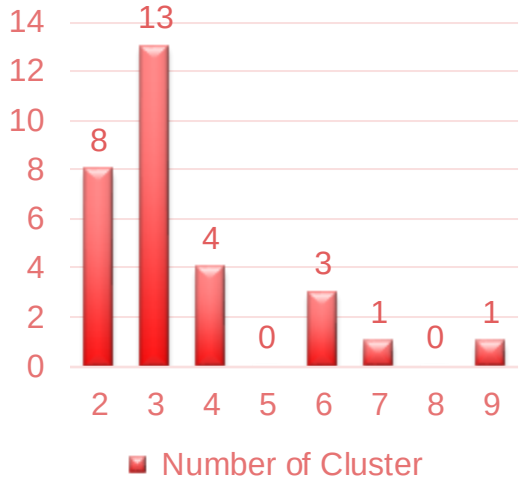
- Stock ticker: AAPL
- Number of matching trends with various tolerances and trend-matching windows
- Observation :
  - Fix the tolerance:
    longer trend-matching window --> fewer matched trends
  - Fix the trend-matching window:
    larger tolerance --> more matched trends
- Conclusion: the trend matching algorithm functions logically

Tolerance/ Days	50%	100%	120%
2 Days	1112	1993	2695
5 Days	5	71	170
10 Days	0	2	9



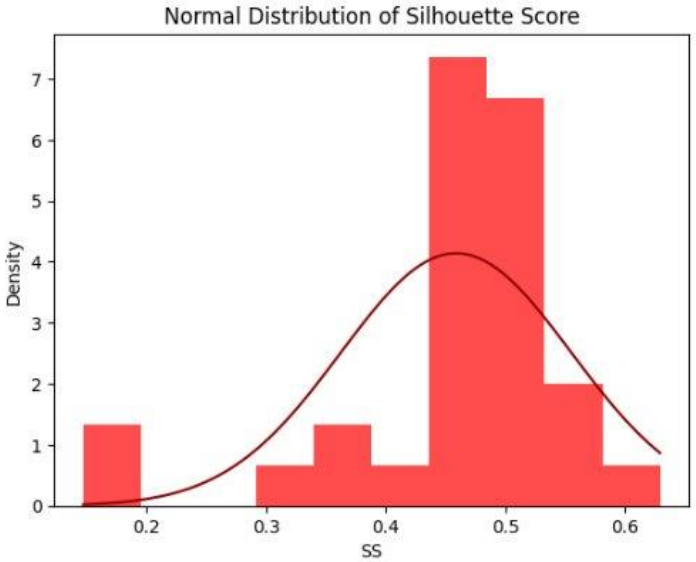
SUBSEQUENT TREND ANALYSIS

- # of clusters
- Randomly picked 30 stocks with
  - tolerance = 100%
  - trend-matching window = 5 days
- 1 cluster is not allowed in the clustering stage
  - The user can see it without clustering
- 2-3 clusters are the most frequent
  - Informative to the user, a small number of main trends
- More than 4 clusters are not often
  - Not so meaningful to the user to determine the Subsequent trend
- Conclusion: 5 days, 100% are a set of good parameters, so we set them as the default



SUBSEQUENT TREND ANALYSIS

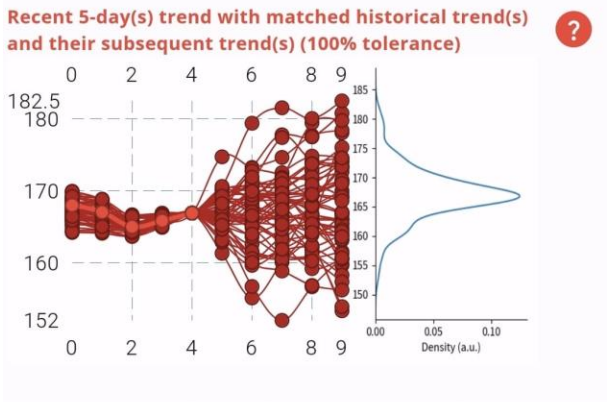
- Silhouette Score
- Randomly picked 30 stocks with
  - tolerance = 100%
  - trend-matching window = 5 days
- Around 0.5 SS are the most frequent
  - well-clustered
  - moderate separation between clusters
- Less than 0.3 SS are not often
  - overlapping clusters or an unclear separation between clusters
- Conclusion: Most of the picked # of clusters are with a reasonable Silhouette Score, so the trends are well-clustered



SUBSEQUENT TREND ANALYSIS

- Stock: AAPL
- Most of Subsequent trend and actual stock price last day return are positive
- The trend of our Subsequent trend analysis in blue dot line is like the actual price
- However, the blue dot line trend is not the most likely Subsequent trend
- Conclusion: Our Subsequent trend analysis can really provide some useful Subsequent trend to user for reference

Subsequent trend analysis on 4/20 - 4/24



Actual stock price on 4/24 - 4/28

VERSION CONTROL

- Powered by Git and GitHub

com4104-grp11 / flutter_application_1 /

Add file

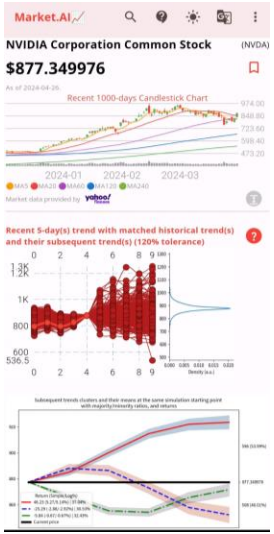
cyruslauwork fixed the issue of sending incorrectly selected last close price — ee3e0a4 · 25 minutes ago — History

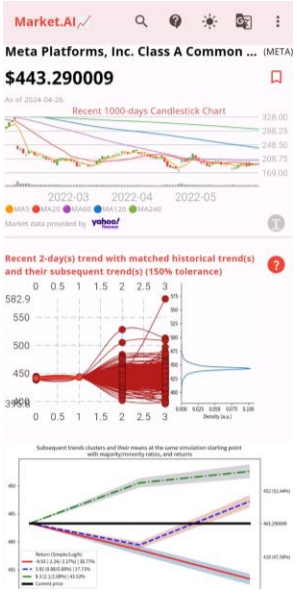
Name	Last commit message	Last commit date
..		
.vscode	add launch.json, hotfix running issues	6 months ago
android	fix Android production network error	1 hour ago
fonts	add lang, sort import name, add fonts	last month
images	fetch listing once, add hsuhk cs dept	last week
ios	change app name, show actions after loading	2 weeks ago
lib	fixed the issue of sending incorrectly selected last close price	25 minutes ago
web	init Flutter app	7 months ago
.gitignore	update README and build proj structure	6 months ago
.metadata	init Flutter app	7 months ago
README.md	add minute candle func	5 months ago
analysis_options.yaml	init Flutter app	7 months ago
pubspec.lock	add appbar linearprogressindicators, update loading animations, save ...	3 days ago
pubspec.yaml	add appbar linearprogressindicators, update loading animations, save ...	3 days ago

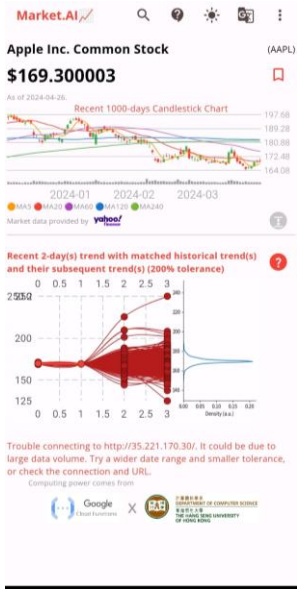
TEST CASES

- Usability
  - Effectiveness ✓
  - Efficiency ✓
  - Satisfaction ✓
  - Error-free ✓
  - Learnability ✓
  - Memorability ✓

- 4 test cases with screen caps

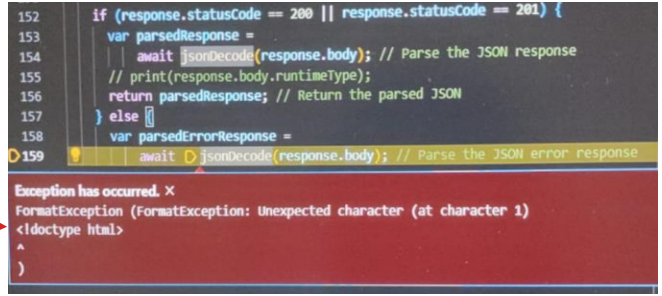


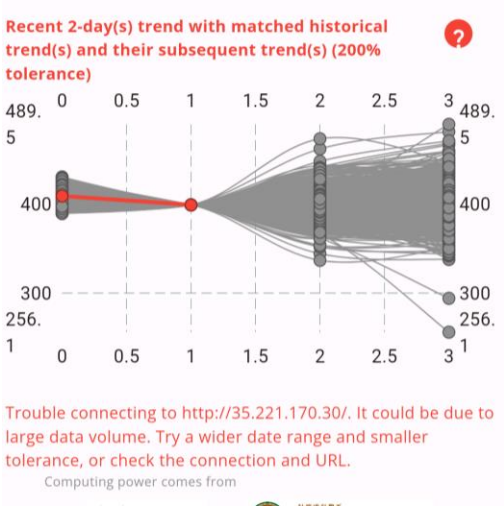




LIMITATION AND FUTURE WORK

- Limitation:
  - Cloud function 'subtrend-to-csv-png' no response when handling large amount of datas (when there are too many matched curves)
  - **Temporary solution:** catch exception, show error message to suggest user to try a wider date range and smaller tolerance
  - **Future:** Use Google Compute Engine VM to handle data processing & have faster processing speed





LIMITATION AND FUTURE WORK

- Future work:
  - Provide extra trend-matching conditions
    - Bollinger band, moving averages over the past 5, 20, 60, 120, and 240 days
  - Search function to support customized start/end dates
  - Use the clustering result for prediction and validate the performance. (hitting chance of the Subsequent trend clusters)
  - Chart pattern scanning
    - Expand the trend matching to more complex patterns

MANPOWER ALLOCATION AND PROGRESS

2024												
Tasks		Assignee(s)	Late Jan	Early Feb	Late Feb	Early Mar	Late Mar	Early Apr	Late Apr	30 Apr		8 May
Functions	Trend match enhancement	Cyrus, Joyce										
	Subsequent trend analytics	Cyrus, Joyce										
	Cloud Compute, Functions	Cyrus, Joyce										
	Gemini Pro API	Cyrus, Joyce										
Improved UI/UX	Trend match	Phoebe, Winnie										
	Subsequent trend analytics	Phoebe, Winnie										
	Final UI/UX	Phoebe, Winnie										
Assessment	Final reports	All								Final Project Presentation		Final Report Submission

REFERENCES

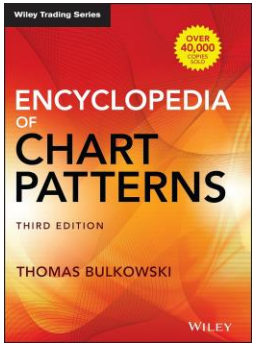
Bibliography

- Bulkowski, T.N., 2021. Encyclopedia of chart patterns. John Wiley & Sons.
- GeeksforGeeks. (2019). *ML | Types of Linkages in Clustering*. [online] Available at: https://www.geeksforgeeks.org/ml-types-of-linkages-in-clustering/.
- PennState: Statistics Online Courses. (n.d.). *14.4 - Agglomerative Hierarchical Clustering | STAT 505*. [online] Available at: https://online.stat.psu.edu/stat505/lesson/14/14.4.
- Preeti, W. (2023). *Stock trading and Investing applications Market size - by platform type (Mobile-based, web-based), operating system (Android, iOS), end user (Professional Traders, individuals), Growth potential, Reginal Analysis & Global Forecast, 2023 - 2032* (No. GMI6691). Global Market Insights Inc.
- Sarkar, T. (2019). *Clustering metrics better than the elbow-method*. [online] Medium. Available at: https://towardsdatascience.com/clustering-metrics-better-than-the-elbow-method-6926e1f723a6.
- Stephanie (2017). *Complete Linkage Clustering*. Statistics How To. Available at: https://www.statisticshowto.com/complete-linkage-clustering/.
- Thakkar, A. and Chaudhari, K., 2021. A comprehensive survey on deep neural networks for stock market: The need, challenges, and future directions. Expert Systems with Applications, 177, p.114800. Available at https://doi.org/10.1016/j.eswa.2021.114800
- 知乎专栏. (n.d.). *肘方法 Elbow of SSE vs 轮廓系数 Silhouette Coefficient*. Available at: https://zhuanlan.zhihu.com/p/668006227

REFERENCES

Thakkar, A. and Chaudhari, K., 2021. A comprehensive survey on deep neural networks for stock market: The need, challenges, and future directions. Expert Systems with Applications, 177, p.114800.

"The stock market has been an attractive field for a large number of organizers and investors to derive useful predictions. Fundamental knowledge of stock market can be utilised with technical indicators to investigate different perspectives of the financial market; also, the influence of various events, financial news, and/or opinions on investors' decisions and hence, market trends have been observed. Such information can be exploited to make reliable predictions and achieve higher profitability."



Bulkowski, T.N., 2021. Encyclopedia of chart patterns. John Wiley & Sons.

"When the smart money trades the securities markets, they leave behind financial footprints. Combine enough footprints together and you have a trail to follow."

ACKNOWLEDGMENT:

We would like to express our sincere gratitude to the Department of Computer Science for the generous sponsorship & support of this project

計算機科學系
DEPARTMENT OF COMPUTER SCIENCE
香港恒生大學
THE HANG SENG UNIVERSITY
OF HONG KONG

06 DEMO