



MẪU BÁO CÁO CỦA MỖI HV

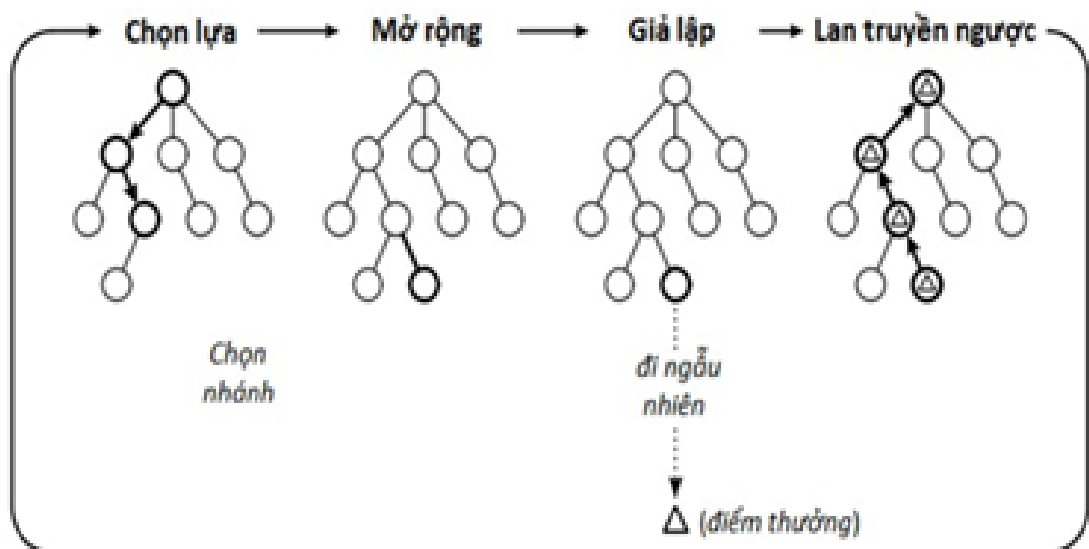
Họ và tên (IN HOA)	HUỖNH THỊ TỔ NGỌC MSHV: CH2002011
Ảnh	
Số buổi vắng	1
Bonus	16
Tên đề tài (VN)	NGHIÊN CỨU, ỨNG DỤNG THUẬT TOÁN MULTITOUR NHẪM TẠO ỨNG DỤNG KHUYẾN NGHỊ ĐA HÀNH TRÌNH CHO KHÁCH DU LỊCH
Tên đề tài (EN)	RESEARCH AND APPLY MULTITOUR ALGORITHM TO CREATE ITINERARY TOURISTS RECOMMENDATION ENGINE
Giới thiệu	<ul style="list-style-type: none"> <i>Bài toán/vấn đề mà đề tài muốn giải quyết</i> Xây dựng một thuật toán tối ưu, chính xác cao để từ đó áp dụng viết một ứng dụng di động có thể đề xuất, gợi ý những hành trình du lịch tối ưu về sự quan tâm, yêu thích, chi phí phù hợp và liên kết lập lịch trình giữa các hành cho một người khi họ muốn đến tham quan các vùng địa lý hoàn toàn mới tại Việt Nam. <i>Lí do chọn đề tài, khả năng ứng dụng thực tế, tính thời sự</i> Tầm quan trọng của du lịch trong thế giới ngày nay là vô cùng to lớn vì nó là một nguồn thu lớn và tạo ra việc làm cho một quốc gia. Khách du lịch phải đối mặt với nhiều thách thức trong quá trình lập kế hoạch hành trình của mình cũng như trong việc lựa chọn các gói tour phù hợp bao gồm

	<p>nhiều hành trình theo sở thích và những ràng buộc khác nhau. Mong muốn với công trình này đề xuất một thuật toán gọi là MULTITOUR nhằm khuyến nghị (đề xuất) cho mọi người những hành trình du lịch phù hợp về sự yêu thích, sự quan tâm của mọi người và tối ưu về chi phí.</p> <ul style="list-style-type: none"> • <i>Mô tả input và output, nên có hình minh họa</i> <p>Input: Sử dụng bộ dữ liệu của Flickr (Flickr Creative Commons 100M - YFCC100M), đây là tập dữ liệu chứa thông tin của các bức ảnh bao gồm ngày/giờ khi bức ảnh được chụp, tọa độ địa lý có độ chính xác cao. Cùng với việc thống kê lượt truy cập người dùng, danh sách các địa điểm tham quan, chi phí truy cập để tạo được bộ dữ liệu đầu vào cho thuật toán.</p> <p>Output: Một ứng dụng di động giúp khách du lịch sẽ được gợi ý những hành trình du lịch mới khi họ muốn thăm quan một vùng đất mới. Thuật toán này sẽ đề xuất những hành trình tối ưu về sự quan tâm, yêu thích, chi phí phù hợp và liên kết lập lịch trình giữa các hành trình đó.</p> 
<p>Mục tiêu</p>	<ul style="list-style-type: none"> • Thu thập và chuyển đổi thông tin từ các hình ảnh trên các website hình ảnh, facebook...v.v để có được tập dữ liệu dồi dào và sát thực về tình hình du lịch tại Việt Nam. • Nghiên cứu, xây dựng thuật toán MULTITOUR để tìm ra logic code, thuật toán có điểm Precision, Recall và F1-Score tốt hơn so với các thuật toán hiện tại. • Áp dụng thuật toán MULTITOUR để xây dựng một ứng dụng di động có thể khuyến nghị các hành trình du lịch phù hợp cho người dân tại Việt Nam.
<p>Nội dung và</p>	<ul style="list-style-type: none"> • Thu thập tập dữ liệu Flickr Creative Commons 100 M (YFCC100M) 100 M Flickr ảnh và video. Sử dụng tập dữ liệu YFCC100M và trích

**phương
pháp thực
hiện**

xuất các hình ảnh được gắn thẻ địa lý được chụp ở các vùng khác nhau trên thế giới. Data thực tế trong đề tài nghiên cứu là tập dữ liệu về danh sách truy cập người dùng, danh sách các POI (các địa điểm du lịch yêu thích), chi phí truy cập của người dùng ở các địa điểm của 8 thành phố năm 2015: (Toronto: Canada, Vienna: Áo, Osaka: Nhật Bản, Budapest: Hungary, Glasgow: Anh, Delhi: Ấn Độ, Edinburgh: Scotland, Perth: Úc)

- Mỗi POI được gắn nhãn với một ID duy nhất và một danh mục như Giải trí, Tôn giáo, v.v. Để đảm bảo độ chính xác của kết quả, các bức ảnh có vị trí địa lý cao nhất độ chính xác đã được chọn. Quá trình tương tự thử nghiệm sử dụng trình tự của người dùng làm cơ sở cho các chuyên thăm thực tế của họ.
- Thuật toán sử dụng chuỗi hành trình địa phương để đo lường các dự đoán của người dùng và chuỗi hành trình của tập dữ liệu toàn cầu để kiểm tra kỹ thuật.
- Để đề xuất nhiều hành trình, thuật toán MULTITOUR đã dùng phương pháp tiếp cận thuật toán Cây tìm kiếm Monte-Carlo (Monte Carlo Tree Search (MCTS)).



Hình 1: Cây tìm kiếm Monte Carlo

	<ul style="list-style-type: none"> • Ví dụ: (các số màu xanh chính là các giá trị quan tâm) <ul style="list-style-type: none"> - I_A: Tôn giáo -> Văn hóa -> Khu bảo tồn -> Tòa nhà - I_B: Giải trí (1.2) -> Công viên (5.6) -> Giải trí (7.8) -> Thể thao (4.4) -> Tòa nhà (3.1) - I_C: Giải trí (2.2) -> Tòa nhà (3.2) -> Khuông viên (3.3) -> Công viên (4.4) -> Tôn giáo (2.2) - I_D: Lịch sử (3.2) -> Tôn Giáo (4.4) -> Văn hóa (5.1) - Người ta tính ra được độ tương đồng giữa hành trình người A và B, C, D theo các chỉ số lần lượt là: 0.338, 0.727, 0.578 - Vì vậy khi người A muốn đi đến Osaka thì sẽ chọn hành trình I_C, I_D vì đây là 2 hành trình có độ tương đồng cao hơn. Từ các POI trong hành trình I_C, I_D ta có thể tạo được nhiều hành trình khác bằng cách thay đổi thứ tự các POI. • Áp bộ dữ liệu vào các thuật toán hiện tại như GNEAR, GPOP, RAND, PERSTOUR, IHA, UBCF-I, TOURINT và TRIPBUILD để kiểm chứng điểm Precision, Recall và F1-Score của thuật toán MULTITOUR là tối ưu hơn. • Xây dựng một ứng dụng di động, sử dụng bộ dữ liệu và thuật toán để khuyến nghị các hành trình du lịch phù hợp cho mọi người.
Kết quả dự kiến	<ul style="list-style-type: none"> • <i>Phần mềm ứng dụng</i>: Dựng được một hệ thống server tự động chuyển đổi thông tin từ các hình ảnh trên các website hình ảnh, facebook...v.v để thu thập thông tin và có được tập dữ liệu dồi dào và sát thực về tình hình du lịch tại Việt Nam. Viết một ứng dụng di động sử dụng thuật toán MULTITOUR và tập dữ liệu thu thập được để khuyến nghị các hành trình du lịch phù hợp cho người dân tại Việt Nam. • <i>Thuật toán MULTITOUR</i>:

```

 $I = \{i_1, i_2, \dots, i_n\}$ 
1  $recommended\_list \leftarrow \emptyset$ 
2  $itinerary\_list \leftarrow \emptyset$   $\triangleright$  list to store itineraries from similar global users
3 for  $user$  in  $matching\_user\_list$  do
4    $itinerary\_list \leftarrow (User, itinerary)$   $\triangleright$  itineraries of similar global users
5 end
6 for  $iterations \leftarrow 1$  to  $max\_loop$  do
7    $I\_list \leftarrow \emptyset$ 
8    $reward\_list \leftarrow \emptyset$ 
9    $list\_copy \leftarrow itinerary\_list$ 
10   $root\_node \leftarrow itinerary$   $\triangleright$  itinerary from the  $itinerary\_list$ 
11   $next\_itinerary \leftarrow itinerary$ 
12   $reward \leftarrow \frac{Int(i)+pop(i)}{f^{cost}(x)}$   $\triangleright$  from Eqn. 25
13   $reward\_list \leftarrow append.reward$ 
14   $count \leftarrow 0$ 
15  while  $count < len(itinerary\_list)$  do
16    if  $next\_itinerary \neq \emptyset$  then
17       $I\_list \leftarrow append.next\_itinerary$ 
18    end

```

- So sánh giữa các phương pháp: Thuật toán MULTITOUR vượt trội hơn hẳn so với các thuật toán cơ bản khác như GNEAR, GPOR, RAND, PERSTOUR, IHA, UBCF-I, TOURINT và TRIPBUILD. Trong bảng 4, 5 và 6 so sánh về điểm Precision, Recall và F1-Score giữa thuật toán MULTITOUR và các thuật toán cơ bản khác trên các tập dữ liệu khác nhau. Từ bảng 4 – 6, hiệu năng tốt nhất và tốt nhì của thuật toán qua mỗi tập dữ liệu được đánh dấu bằng in đậm và gạch chân.

Table 4
Comparison between MULTITOUR and different baseline algorithms in terms of Precision using different datasets.

Algorithms	MULTITOUR	IHA	PERSTOUR	TOURINT	TRIPBUILD	UBCF	GPOR	GNEAR	RAND
Delhi-Edinburgh	0.395 \pm 0.010 3	0.500 \pm 0.027 1	0.333 \pm 0.043 4.5	0.333 \pm 0.013 4.5	0.308 \pm 0.037 6	0.417 \pm 0.018 2	0.281 \pm 0.029 8	0.286 \pm 0.014 7	0.241 \pm 0.025 9
Osaka-Edinburgh	0.500 \pm 0.022 1	0.281 \pm 0.021 6	0.348 \pm 0.033 4	0.429 \pm 0.018 2	0.250 \pm 0.041 7	0.389 \pm 0.036 3	0.313 \pm 0.008 5	0.211 \pm 0.03 8	0.188 \pm 0.019 9
Vienna-Edinburgh	0.476 \pm 0.015 3	0.363 \pm 0.031 2	0.607 \pm 0.043 1	0.375 \pm 0.034 5	0.267 \pm 0.02 9	0.333 \pm 0.013 6.5	0.455 \pm 0.016 4	0.296 \pm 0.041 8	0.333 \pm 0.026 6.5
Delhi-Osaka	0.609 \pm 0.017 1	0.455 \pm 0.028 3	0.512 \pm 0.013 2	0.412 \pm 0.027 4	0.381 \pm 0.035 5	0.343 \pm 0.029 6	0.294 \pm 0.021 7	0.245 \pm 0.044 9	0.273 \pm 0.047 8
Glasgow-Edinburgh	0.239 \pm 0.036 3	0.333 \pm 0.023 1	0.226 \pm 0.025 4	0.267 \pm 0.014 2	0.196 \pm 0.024 6	0.176 \pm 0.033 7	0.209 \pm 0.028 5	0.167 \pm 0.011 8.5	0.167 \pm 0.021 8.5
Delhi-Budapest	0.500 \pm 0.043 1.5	0.478 \pm 0.041 3	0.500 \pm 0.017 1.5	0.293 \pm 0.032 9	0.438 \pm 0.019 4	0.409 \pm 0.024 5	0.357 \pm 0.016 7	0.381 \pm 0.013 6	0.316 \pm 0.014 8
Budapest-Edinburgh	0.340 \pm 0.022 4	0.481 \pm 0.015 1	0.256 \pm 0.009 6	0.228 \pm 0.014 8	0.391 \pm 0.032 2	0.216 \pm 0.044 9	0.316 \pm 0.023 5	0.258 \pm 0.044 7	0.337 \pm 0.017 3
Delhi-Vienna	0.529 \pm 0.044 1	0.326 \pm 0.027 4	0.292 \pm 0.035 6	0.313 \pm 0.028 5	0.332 \pm 0.011 9	0.361 \pm 0.021 7	0.417 \pm 0.017 3	0.235 \pm 0.018 8	0.364 \pm 0.038 3
Delhi-Glasgow	0.500 \pm 0.009 2.5	0.448 \pm 0.018 5	0.643 \pm 0.042 1	0.500 \pm 0.027 2.5	0.406 \pm 0.035 7	0.381 \pm 0.029 8	0.421 \pm 0.021 6	0.476 \pm 0.044 4	0.366 \pm 0.047 9
Budapest-Toronto	0.528 \pm 0.029 2	0.393 \pm 0.015 7	0.632 \pm 0.031 1	0.435 \pm 0.042 5	0.474 \pm 0.011 4	0.341 \pm 0.016 9	0.421 \pm 0.038 6	0.357 \pm 0.022 8	0.481 \pm 0.017 3
Budapest-Vienna	0.447 \pm 0.021 3	0.412 \pm 0.015 4	0.56 \pm 0.043 1	0.309 \pm 0.038 7	0.333 \pm 0.018 6	0.471 \pm 0.045 2	0.255 \pm 0.019 9	0.278 \pm 0.028 8	0.367 \pm 0.024 5
Budapest-Glasgow	0.550 \pm 0.017 1	0.379 \pm 0.018 4	0.400 \pm 0.038 2.5	0.4 \pm 0.03 2.5	0.236 \pm 0.038 9	0.296 \pm 0.01 7	0.353 \pm 0.017 5	0.323 \pm 0.018 6	0.205 \pm 0.038 8

Table 5
Comparison between MULTITOUR and different baseline algorithms in terms of Recall using different datasets.

Algorithms	MULTITOUR	IHA	PERSTOUR	TOURINT	TRIPBUILD	UBCF	GPOR	GNEAR	RAND
Delhi-Edinburgh	0.300 \pm 0.032 1	0.24 \pm 0.015 2	0.200 \pm 0.009 4	0.220 \pm 0.016 3	0.160 \pm 0.042 5.5	0.100 \pm 0.012 9	0.120 \pm 0.034 8	0.160 \pm 0.026 5.5	0.140 \pm 0.02 7
Osaka-Edinburgh	0.300 \pm 0.026 1.5	0.225 \pm 0.017 3	0.200 \pm 0.009 4	0.300 \pm 0.013 1.5	0.15 \pm 0.012 6	0.175 \pm 0.034 5	0.125 \pm 0.015 7	0.100 \pm 0.01 8	0.075 \pm 0.027 9
Vienna-Edinburgh	0.286 \pm 0.032 3	0.257 \pm 0.014 4	0.400 \pm 0.028 1	0.171 \pm 0.044 7	0.114 \pm 0.023 9	0.314 \pm 0.025 2	0.143 \pm 0.038 8	0.229 \pm 0.022 5	0.200 \pm 0.017 6
Delhi-Osaka	0.333 \pm 0.016 2	0.238 \pm 0.036 5	0.500 \pm 0.039 1	0.167 \pm 0.048 7	0.190 \pm 0.043 6	0.286 \pm 0.02 4	0.119 \pm 0.049 9	0.310 \pm 0.034 3	0.143 \pm 0.045 8
Glasgow-Edinburgh	0.196 \pm 0.022 3	0.250 \pm 0.009 1	0.214 \pm 0.032 2	0.143 \pm 0.017 6.5	0.179 \pm 0.018 4	0.107 \pm 0.038 8	0.161 \pm 0.044 5	0.143 \pm 0.027 6.5	0.089 \pm 0.019 9
Delhi-Budapest	0.337 \pm 0.033 1	0.224 \pm 0.021 3	0.204 \pm 0.045 4	0.245 \pm 0.03 3	0.143 \pm 0.038 7	0.184 \pm 0.01 5	0.102 \pm 0.026 9	0.163 \pm 0.046 6	0.122 \pm 0.047 8
Budapest-Edinburgh	0.362 \pm 0.028 1	0.277 \pm 0.021 2	0.213 \pm 0.027 4	0.149 \pm 0.018 7	0.191 \pm 0.039 5	0.234 \pm 0.019 3	0.128 \pm 0.031 8	0.170 \pm 0.04 6	0.106 \pm 0.016 9
Delhi-Vienna	0.231 \pm 0.021 4	0.385 \pm 0.031 1	0.179 \pm 0.016 6	0.128 \pm 0.039 8	0.282 \pm 0.02 2	0.154 \pm 0.019 7	0.256 \pm 0.014 3	0.103 \pm 0.024 9	0.205 \pm 0.033 5
Delhi-Glasgow	0.455 \pm 0.019 1	0.295 \pm 0.021 4.5	0.205 \pm 0.038 8	0.25 \pm 0.017 6	0.285 \pm 0.028 4.5	0.364 \pm 0.013 2	0.182 \pm 0.016 9	0.227 \pm 0.036 7	0.341 \pm 0.039 3
Budapest-Toronto	0.514 \pm 0.041 1	0.297 \pm 0.026 6	0.324 \pm 0.048 5	0.27 \pm 0.044 7	0.243 \pm 0.023 8	0.378 \pm 0.025 3	0.216 \pm 0.047 9	0.405 \pm 0.033 2	0.351 \pm 0.024 4
Budapest-Vienna	0.395 \pm 0.026 2.5	0.488 \pm 0.05 1	0.326 \pm 0.027 5	0.395 \pm 0.042 2.5	0.349 \pm 0.011 4	0.186 \pm 0.031 9	0.362 \pm 0.021 6	0.233 \pm 0.013 8	0.256 \pm 0.032 7
Budapest-Glasgow	0.268 \pm 0.028 4	0.415 \pm 0.011 1	0.293 \pm 0.021 3	0.195 \pm 0.044 7.5	0.317 \pm 0.027 2	0.3195 \pm 0.035 7.5	0.146 \pm 0.028 9	0.244 \pm 0.031 5	0.220 \pm 0.021 6

Algorithms	MULTITOUR	IHA	PERSTOUR	TOURINT	TRIPBUILD	UBCF	GPOF	GNEAR	RAND
Delhi-Edinburgh	0.341 ± 0.017 1	0.324 ± 0.033 2	0.290 ± 0.028 4	0.265 ± 0.011 3	0.211 ± 0.008 5	0.161 ± 0.036 9	0.164 ± 0.019 8	0.205 ± 0.007 6	0.177 ± 0.022 7
Osaka-Edinburgh	0.375 ± 0.035 1	0.250 ± 0.038 4	0.254 ± 0.014 3	0.353 ± 0.096 2	0.188 ± 0.024 6	0.241 ± 0.032 5	0.179 ± 0.029 7	0.136 ± 0.007 8	0.107 ± 0.011 9
Vienna-Edinburgh	0.357 ± 0.033 2	0.353 ± 0.024 3	0.500 ± 0.011 1	0.235 ± 0.022 7	0.16 ± 0.029 9	0.324 ± 0.042 4	0.217 ± 0.048 8	0.258 ± 0.047 5	0.250 ± 0.046 6
Delhi-Osaka	0.431 ± 0.014 2	0.313 ± 0.026 3	0.506 ± 0.041 1	0.237 ± 0.019 7	0.254 ± 0.025 6	0.312 ± 0.038 4	0.169 ± 0.009 9	0.274 ± 0.02 5	0.188 ± 0.035 8
Glasgow-Edinburgh	0.216 ± 0.016 3	0.286 ± 0.039 1	0.220 ± 0.02 2	0.186 ± 0.037 5	0.187 ± 0.019 4	0.133 ± 0.029 8	0.182 ± 0.023 6	0.154 ± 0.034 7	0.116 ± 0.05 9
Delhi-Budapest	0.395 ± 0.014 1	0.306 ± 0.039 2	0.29 ± 0.036 3	0.267 ± 0.011 4	0.215 ± 0.044 7	0.254 ± 0.036 5	0.159 ± 0.012 9	0.229 ± 0.042 6	0.176 ± 0.038 8
Budapest-Edinburgh	0.351 ± 0.049 1.5	0.351 ± 0.032 1.5	0.244 ± 0.041 4	0.179 ± 0.038 8	0.257 ± 0.026 3	0.224 ± 0.05 5	0.182 ± 0.034 7	0.205 ± 0.043 6	0.164 ± 0.042 9
Delhi-Vienna	0.321 ± 0.023 2	0.353 ± 0.034 1	0.222 ± 0.05 6	0.182 ± 0.036 8	0.242 ± 0.026 5	0.194 ± 0.025 7	0.317 ± 0.037 3	0.143 ± 0.019 9	0.262 ± 0.029 4
Delhi-Glasgow	0.476 ± 0.048 1	0.356 ± 0.043 3	0.31 ± 0.02 7	0.333 ± 0.016 6	0.342 ± 0.026 5	0.372 ± 0.041 2	0.254 ± 0.049 9	0.308 ± 0.034 8	0.353 ± 0.045 4
Budapest-Toronto	0.521 ± 0.032 1	0.338 ± 0.034 6	0.429 ± 0.02 2	0.333 ± 0.013 7	0.321 ± 0.043 8	0.359 ± 0.04 5	0.286 ± 0.014 9	0.38 ± 0.028 4	0.406 ± 0.046 3
Budapest-Vienna	0.420 ± 0.039 2	0.447 ± 0.022 1	0.412 ± 0.016 3	0.347 ± 0.048 4	0.341 ± 0.02 5	0.267 ± 0.037 8	0.277 ± 0.026 7	0.253 ± 0.03 9	0.301 ± 0.034 6
Budapest-Glasgow	0.361 ± 0.026 2	0.391 ± 0.046 1	0.338 ± 0.047 3	0.262 ± 0.044 6	0.271 ± 0.027 5	0.235 ± 0.035 8	0.257 ± 0.033 9	0.278 ± 0.021 4	0.240 ± 0.045 7

• *Bộ dữ liệu, etc*

- Dữ liệu truy cập của người dùng (userVisits-ijcai15): Tập dữ liệu này bao gồm một tập hợp người dùng và lượt truy cập của họ đến các điểm ưa thích (POI) khác nhau ở tám thành phố. Lượt truy cập POI của người dùng được xác định dựa trên ảnh Flickr YFCC100M được gắn thẻ địa lý: (i) được ánh xạ tới vị trí POI cụ thể và các danh mục POI; và (ii) được nhóm thành các chuỗi du lịch riêng lẻ (lượt truy cập POI của người dùng liên tiếp chênh lệch nhau <8 giờ).
- Danh sách POI (poiList-ijcai15): Tập dữ liệu này bao gồm các điểm ưa thích (POI) khác nhau được tìm thấy ở mỗi thành phố trong số tám thành phố, dựa trên các mục nhập/bài viết của họ trên Wikipedia. Thông tin như tên POI, tọa độ vĩ độ / vĩ độ và chủ đề (danh mục) cũng được lấy từ Wikipedia.
- Bảng chi phí-lợi nhuận POI (costProf-ijcai15): Tập dữ liệu này bao gồm các bảng chi phí-lợi nhuận khác nhau (cho tám thành phố) cho biết chi phí (dựa trên khoảng cách) cần thiết để đi từ một điểm ưa thích (POI) đến một điểm yêu thích khác và lợi nhuận kết quả (dựa trên mức độ phổ biến) thu được từ đạt POI đó.
- **Link dataset:**
<https://sites.google.com/site/limkwanhui/datacode?authuser=0>

Tài liệu tham khảo

- [1]. Joy Lal Sarkara, Abhishek Majumdera, Chhabi Rani Panigrahib, Sudipta Royc. "MULTITOUR: A multiple itinerary tourists recommendation engine". In Electronic Commerce Research and Applications 40 (2020) 100943.
- [2]. Kwan Hui Lim, Jeffrey Chan, Christopher Leckie and Shanika Karunasekera. "Personalized Tour Recommendation based on User Interests and Points of Interest Visit Durations". In Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15). Pg 1778-1784. Jul 2015.
- [3]. Kwan Hui Lim, Jeffrey Chan, Christopher Leckie and Shanika Karunasekera. "Towards Next Generation Touring: Personalized Group Tours". In Proceedings of the 26th International Conference on Automated Planning and Scheduling (ICAPS'16). Pg 412-420. Jun 2016.

