

Rating Inference for Custom Trips from Enriched GPS Traces using Random Forests

Theodoros Chondrogiannis

University of Konstanz

Germany

theodoros.chondrogiannis@uni-konstanz.de

Mouzhi Ge

Deggendorf Institute of Technology

Germany

mouzhi.ge@th-deg.de

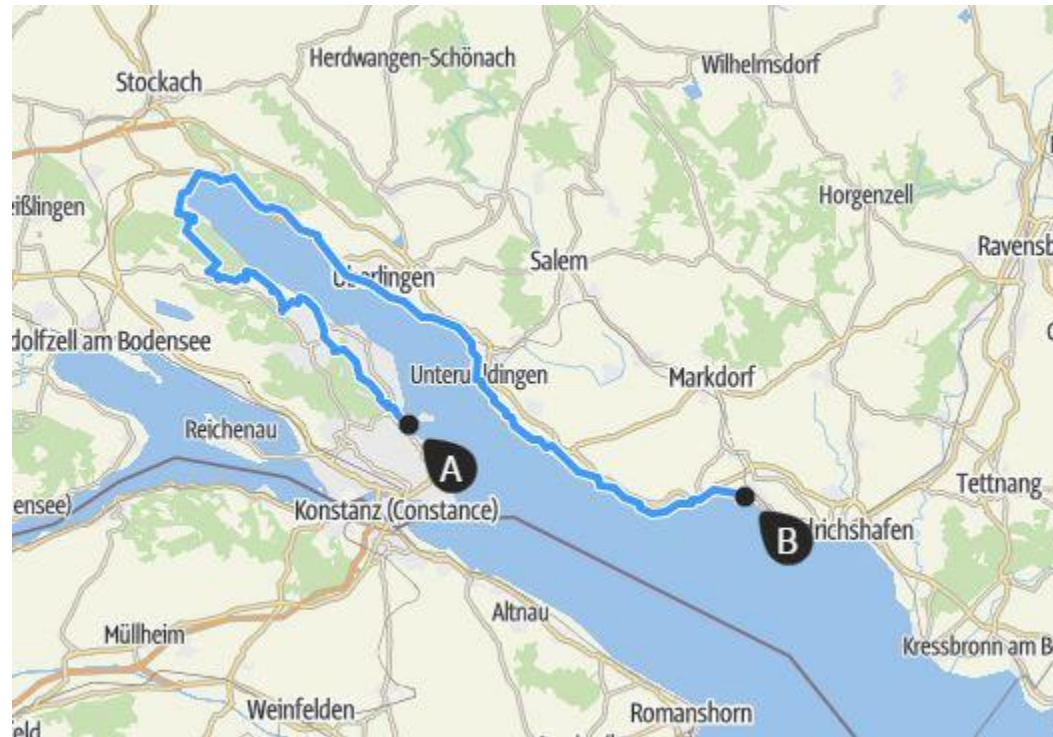
Agenda

- Motivation
- Definition and Background
- Problem statement and scope
- Trip rating inference framework
- Feature encoders
- Experimental settings and results
- Key take-aways

Motivation

- **GPS-enabled devices** allow us to pinpoint our location and generate a large amount of data that traces our movements along trips.
- **Custom trips** are designed to cater to travelers' specific desires and user preferences for personalized tourism experience.
- Since the custom trip is usually new in the system, **no rating** can be shown to the user. As a result, the **rating inference** of custom trips has emerged as an important feature in tourism applications and location-based services.
- This paper aims to determine **which representation feeds best to the machine learning algorithms** and achieves higher accuracy for rating inference.
- Apart from **trip recommendations**, rating inference in this paper can be considered a **second opinion for custom trips** defined by users.

Custom trip

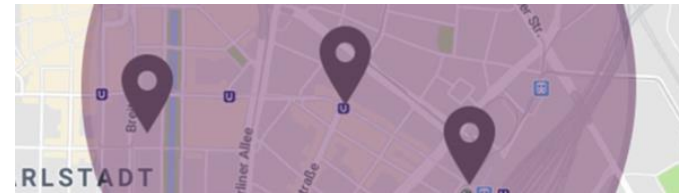


Enriched GPS traces along with the trip

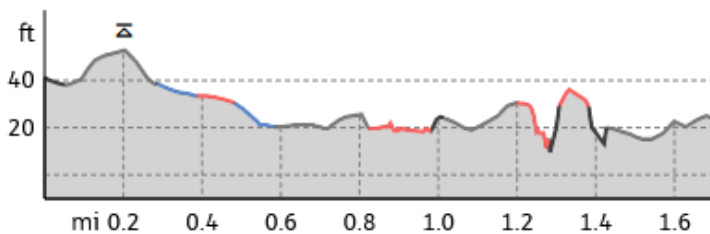
- Trip location



- Closeness to POIs



- Trip elevation



- Closeness to places where users take pictures



Multi-criteria ratings for trips

- Multi-criteria ratings consider different factors simultaneously. For example, one hiking route would have various attributes for ratings, such as Condition, Difficulty, Technique, Quality of Experience, and Landscape.

Condition	★	★	★	☆	☆
Difficulty	★	★	☆	☆	☆
Technique	★	★	★	★	☆
Quality of Experience	★	★	★	★	☆
Landscape	★	★	★	☆	☆

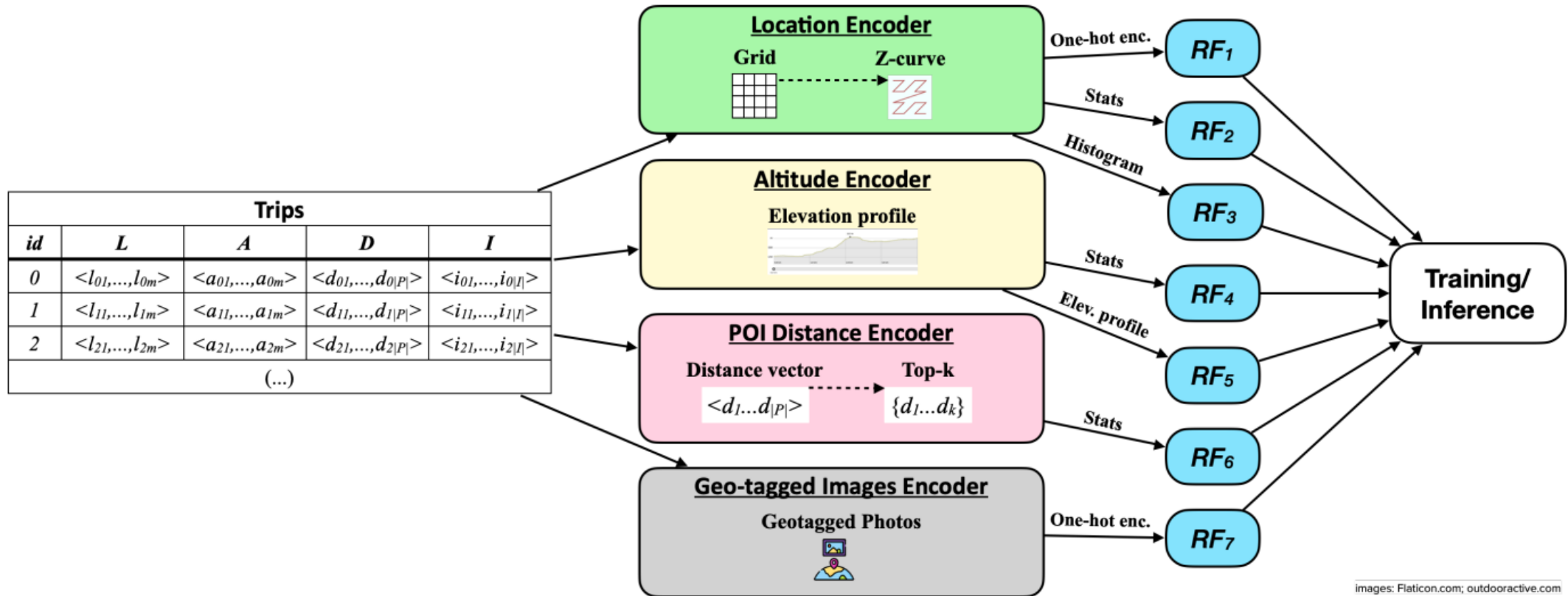
Problem statement and scope

- The user designs a custom trip
- This trip contains enriched GPS traces
- There are different rating criteria for this trip

- We want to infer/predict
- The rating of each criterion for this custom trip

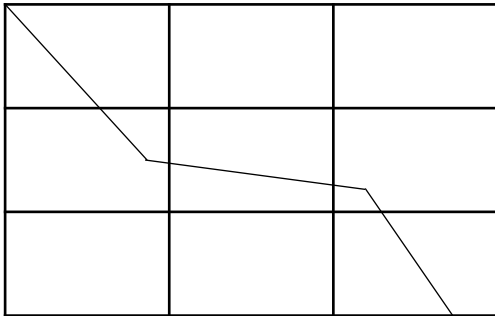
- We want to use machine learning to do the rating inference, but the focus of this work is not ML model selection, it is feature engineering and encoding selections, given the enriched GPS traces are complex.

Trip rating inference framework

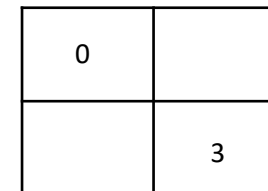
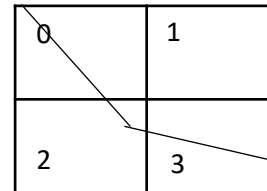
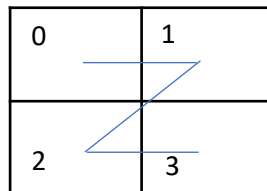


Location Encoder

- We first impose a $n \times n$ grid over the space defined by the minimum bounding rectangle of all traces.



- One-hot encoding
- Z-order curve to first ID the grids
 - For each set of IDs the trip crosses, a vector that contains basic statistics, i.e., min, max, mean, and median values.
 - Histogram of n buckets



Altitude Encoder

<

- total ascent,
- total descent
- minimum altitude
- maximum altitude
- standard deviation of the elevation profile

>

POI Distance Encoder

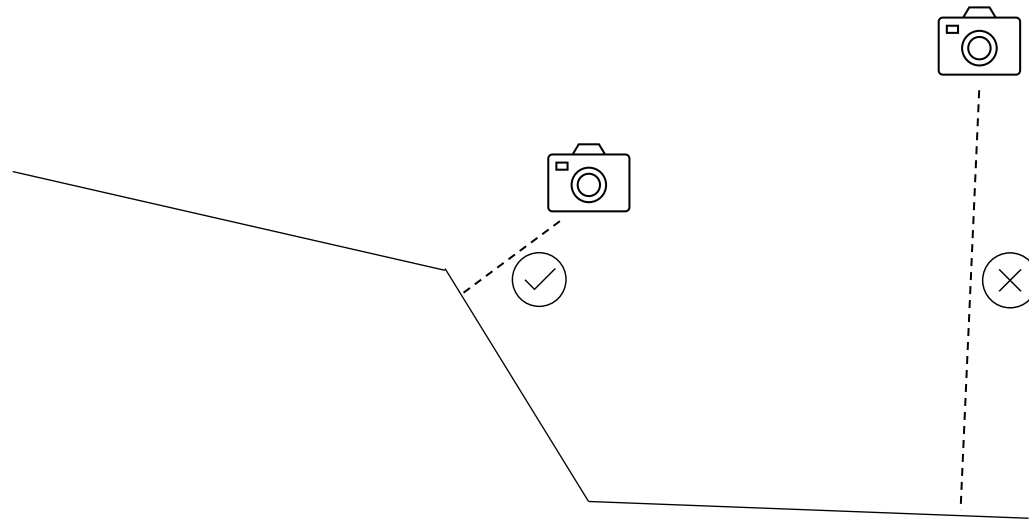
- A combination of two vectors
- Vector 1: Distances to all POIs

Plus

- Vector 2: Distances to a predefined set of POIs (k nearest POIs)

Geo-tagged Images Encoder

- A bit vector and the size equals to the number of images in I
- we set each bit associated with an image to 1 if the minimum distance between the trace and the image location is below a predefined threshold, e.g., 20 meters.



Datasets

- Trip data obtained from Outdooractive: www.outdooractive.com

network	trips	avg # points	avg length (km)
Hiking	1,813	496.05	12.207
Cycling	4,051	1489.86	49.617

- Elevation data for trip from Copernicus: www.copernicus.eu
- 181,185 POI data from www.kaggle.com/datasets/ehallmar/points-of-interest-poi-database
- 50,000 geotagged images from www.kaggle.com/datasets/habedi/large-dataset-of-geotagged-images

Encoding methods overview

One-hot Enc. Location encoder that imposes a grid and uses one-hot encoding to indicate which cells are crossed by the trip trace.

Grid Stats Location encoder that imposes a grid, determines the Z-order of the grid cells, and computes basic statistics of the numerical cell IDs.

Histogram Location encoder that imposes a grid, determines the Z-order of the grid cells, and computed a histogram.

Elevation Altitude encoder using statistics computed over the elevation profile.

POIs POI Distance Encoder that computes a vector of distances to a predefined set of POIs.

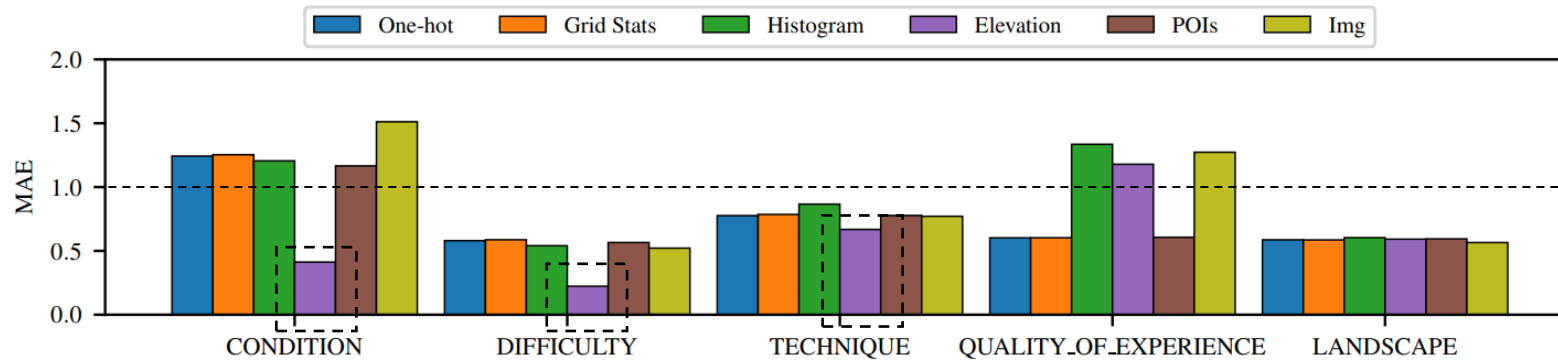
Img Geo-tagged Images Encoder that computes distances to a predefined set of geotagged images and uses one-hot encoding to indicate which images have been taken from location alongside the trip trace.

ML Model and Evaluation Metric

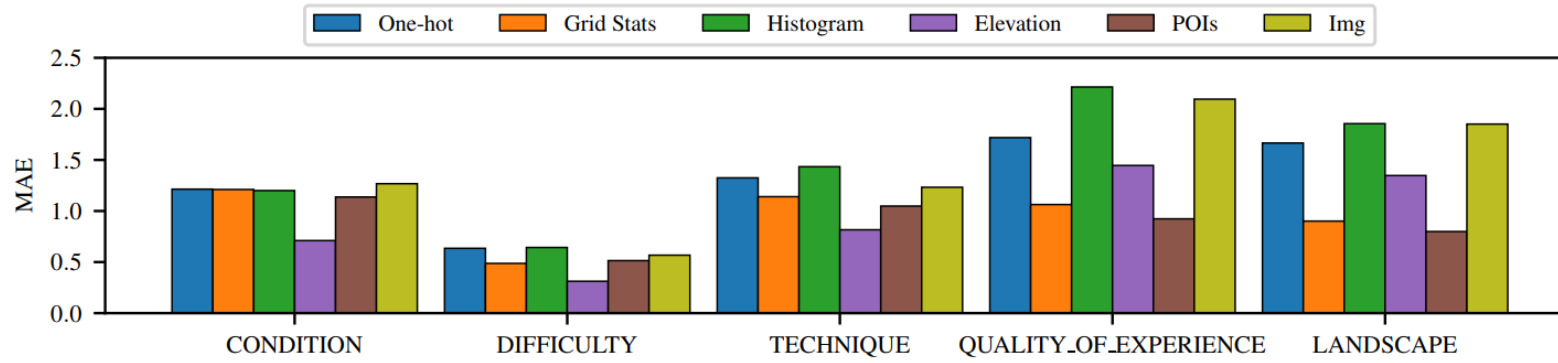
- Random forest
 - Our previous experiments demonstrated that Random Forest performs best in several similar rating inference scenarios. We used Random Forest classifier in this work.
- MAE, widely used for evaluating rating predictions, especially in recommender system research.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

Experimental Results



(a) Hiking data set



(b) Cycling data set

Take aways

- "One size fits all" encodings may lower the quality of multi-criteria rating inferences.
- Different encodings might be dynamically used to infer different rating criteria.
- The trip-oriented ratings are focused on the intrinsic features of the trip. Thus, the encoding of trip profiles can offer higher-quality rating inferences.
- User-oriented ratings focus on how users feel about the trip and user satisfaction.

Summary and Future Research

- Scope of this research: **encoding selection, not model selection.**
- The model may consider more **contextual factors**. For example, the context of a trip may include group dynamics, previous experiences, and cultural factors.
- Users would often like to know how the inference is made. In turn, users can be more confident in their trip decisions. Therefore, developing **transparent and explainable** models may increase user trust and satisfaction.
- Including **user feedback** to enhance user engagement is also critical. User feedback can be used to improve the model training and provide continuous improvement for implementing trip recommendations

