Inferring Socioeconomic Characteristics from One-Week Activity Sequences Using TabTransformer

Sungho Lim
Civil and Environmental Engineering Department
University of Michigan-Ann Arbor
Ann Arbor, USA
lsungho@umich.edu

***Abstract*—This study explores the feasibility of inferring individual-level socioeconomic attributes—age group, education level, and household income—from weekly activity sequences. Using data from the 2018 Bay Area Travel Survey (BATS), we construct 7-day sequences of activity stays with encoded temporal, spatial, and semantic features. A TabTransformer model is employed to jointly predict the target attributes from each individual's activity pattern. The model processes both categorical and continuous features through embedding and projection layers, followed by a shared Transformer encoder and separate classification heads. Results show the highest accuracy for education level (81.8%), followed by household income (51.6%) and age group (48.3%). These findings demonstrate the potential of behavioral sequence modeling to enrich survey-based population synthesis with inferred socioeconomic information.**

***Index Terms*—Socioeconomic inference, activity sequences, household travel survey, multi-task learning, Transformer model**

I. Introduction

Understanding the relationship between individuals' socioeconomic characteristics and their daily activity-travel patterns is a central element in travel demand forecasting. Forecasting models often rely on assumptions about future socioeconomic profiles—such as age, income, and household composition—to estimate how travel demand will evolve over time. Thus, learning and quantifying how socioeconomic status influences activity patterns is essential not only for analyzing current behavior but also for predicting future mobility needs.

Moreover, this understanding provides a critical bridge between traditional survey-based data and modern passively collected data sources. Household Travel Surveys (HTS) provide rich, high-quality information about individual behaviors and attributes, but they are costly and infrequent. If we can learn robust relationships between activity patterns and demographics from HTS data, we can apply those models to passive data sources like location-based services (LBS), enabling near real-time behavioral monitoring without waiting for the next survey cycle.

Recent studies have explored data-driven approaches to infer demographic attributes from mobility behavior. For instance, Zhang et al. [1] used similarity measures and SVMs for demographic inference from multi-dimensional activity patterns. Wang et al. [2] introduced DeepSEI, combining deep and recurrent networks. Other works, including Wu et al. [3] and Kim et al. [4], emphasized the value of contextual features and generative modeling techniques. Collectively, these studies demonstrate the growing feasibility of learning demographic information from structured behavioral data.

With continued advances in machine learning, Transformer-based models such as the TabTransformer [5] offer new opportunities to learn from sequences of structured categorical data. This study uses the TabTransformer to predict a person's age group, education level, and household income based on one week of activity sequences. While the model is trained on HTS data from the Bay Area Travel Survey (BATS), the broader aim is to create a transferable framework that can be applied to passively collected mobility data.

II. Method

A. Problem Formulation

The goal is to predict three socioeconomic attributes—**age group**, **education level**, and **household income**—from weekly activity patterns. The input is a **7-day sequence of activity stays**, constructed by converting trips into stay episodes anchored to specific locations. Compared to traditional trip-based origin-destination formats, this representation preserves richer temporal and semantic context.

Each stay is encoded using categorical features, including the activity type (e.g., home, work, shopping), day of the week, start time bin, duration bin, and the stay's sequence position within the week. These attributes serve as inputs to a **multi-task classification model** that jointly predicts the target demographic variables.

B. Dataset Description

We use data from the **2018 Bay Area Travel Survey (BATS)**, conducted by the Metropolitan Transportation Commission (MTC). The dataset includes:

- Approximately 5,000 individuals
- Around 190,000 trips recorded over 7 days
- Detailed attributes on trip purpose, timing, locations, and respondent demographics

TABLE I
GROUPED CATEGORIES FOR DEMOGRAPHIC VARIABLES

Variable	Group	Raw Categories
Income	Low income	Under $25k, $25k–$49,999
	Mid income	$50k–$74,999, $75k–$99,999, $100k–$249,999
	High income	$250k or more
Age	Child	Under 5, 5–15, 16–17
	Adult	18–24, 25–34, 35–44, 45–54, 55–64
	Senior	65–74, 75 or older
Education	HS or less	Less than high school, High school graduate/GED
	Some college	Some college, Vocational/technical training, Associate degree
	Bachelor+	Bachelor's degree, Graduate/Post-graduate degree

C. Data Processing Pipeline

Raw travel survey records are transformed into structured activity sequences through several preprocessing steps. Entries with invalid or missing income and education responses are removed, and demographic attributes (age, education, income) are grouped into broader categories to reduce class imbalance and improve generalizability.

Temporal features—such as duration, day of the week, and time of day—are extracted from activity timestamps. Sine and cosine transformations are applied to encode daily and weekly cycles, and each activity is assigned a relative day index for alignment within the 7-day window.

Categorical features are label-encoded, and continuous features are standardized. For each individual, up to 50 activities are retained and padded as needed. The dataset is then split into training and validation sets and loaded in mini-batches using a custom PyTorch Dataset.

D. Model Architecture and Training

We implement a Transformer-based sequence classifier to jointly predict age group, education level, and income group. Each activity is represented by categorical features (embedded and summed) and continuous features (projected via a linear layer), which are combined into a unified activity embedding.

A classification token is prepended to the sequence, and positional encodings are added. The sequence is processed by a multi-layer Transformer encoder, and the final output corresponding to the classification token is passed to three separate heads for multi-task prediction.

The model is trained using the Adam optimizer (learning rate = 0.001) to minimize the sum of cross-entropy losses. Automatic mixed precision (AMP) is used to improve training efficiency, and accuracy is reported separately for each target on the validation set.

III. RESULTS

Among the three targets, education was predicted most accurately (81.8%), primarily due to strong performance for the bachelor's degree or higher group (F1: 0.899), while the model failed to identify individuals with a high school education or less (F1: 0.000). Age prediction reached moderate accuracy (48.3%), with adults classified best (F1: 0.564) and poor performance for seniors (F1: 0.168). Income prediction was less successful (51.6% accuracy), performing reasonably for mid-income individuals (F1: 0.674), but poorly for low-income (F1: 0.283) and high-income groups (F1: 0.000).

TABLE II
CLASSIFICATION ACCURACY BY CLASS AND OVERALL (AGE, EDUCATION, INCOME)

Class	Age	Education	Income
Class 0	0.452 (Child)	0.0000 (HS or less)	0.497 (Low)
Class 1	0.564 (Adult)	0.024 (Some college)	0.652 (Mid)
Class 2	0.147 (Senior)	0.978 (Bachelor+)	0.000 (High)
Accuracy	0.483	0.818	0.516

TABLE III
CLASSIFICATION PERFORMANCE BY TARGET AND CLASS

Target	Class	Precision	Recall	F1-score
Age	Child (0)	0.420	0.451	0.435
	Adult (1)	0.566	0.563	0.564
	Senior (2)	0.197	0.146	0.168
Education	HS or less (0)	0.000	0.000	0.000
	Some college (1)	0.146	0.023	0.040
	Bachelor+ (2)	0.833	0.977	0.899
Income	Low (0)	0.198	0.497	0.283
	Mid (1)	0.697	0.652	0.674
	High (2)	0.000	0.000	0.000

IV. CONCLUSION

This study demonstrates the feasibility of inferring key demographic attributes—such as age group, and income—from 7-day activity sequences using various predictive models. Our results highlight the relative strengths of different model classes and reveal which behavioral features are most informative for each target. These findings suggest that activity-based behavior alone can offer meaningful signals for socioeconomic inference.

While our analysis is limited to a single region and one week of data, the framework is adaptable to other settings and can complement traditional survey methods. Future work will extend this approach to generative models to support synthetic population generation for simulation-based applications.

REFERENCES

[1] B. Zhang, S. Rasouli, and T. Feng, "Social demographics imputation based on similarity in multi-dimensional activity-travel pattern: A two-step approach," *Travel Behav. Soc.*, vol. 37, p. 100843, 2024, doi: 10.1016/j.tbs.2024.100843.

[2] Z. Wang, M. Liu, C. Long, Q. Zhang, J. Li, and C. Miao, "On inferring user socioeconomic status with mobility records," in *Proc. IEEE Int. Conf. Big Data*, Osaka, Japan, 2022, pp. 646–655, doi: 10.1109/BigData55660.2022.10020870.

[3] L. Wu, L. Yang, Z. Huang, Y. Wang, Y. Chai, X. Peng, and Y. Liu, "Inferring demographics from human trajectories and geographical context," *Comput. Environ. Urban Syst.*, vol. 77, p. 101368, 2019, doi: 10.1016/j.compenvurbsys.2019.101368.

[4] E.-J. Kim, D.-K. Kim, and K. Sohn, "Imputing qualitative attributes for trip chains extracted from smart card data using a conditional generative adversarial network," *Transp. Res. C Emerg. Technol.*, vol. 137, p. 103616, 2022, doi: 10.1016/j.trc.2022.103616.

[5] X. Huang, A. Khetan, M. Cvitkovic, and Z. Karnin, "TabTransformer: Tabular data modeling using contextual embeddings," *arXiv preprint arXiv:2012.06678*, 2020.