



Lexicon for Volume and Turning Movements Project (November 22, 2017)

The Lexicon is intended to complement other project deliverables by providing additional detail and background. In order to facilitate the broadest possible audience, the authors have incorporated a glossary of relevant terms which is the primary content of the Lexicon.

A

AADT (Acronym) • Annual Average Daily Traffic

ANN (Analysis) • Artificial Neural Network – a machine learning model

API (Acronym) • Application Programming Interface – facilitates queries

Archive Data (Data) • Historical data

ATIS (Acronym) • Advanced Traveler Information System

ATR Station (Acronym) • Automatic Traffic Recorder Station – traffic sensor

C

Calibrated Count Stations (Data) • Stations providing accurate vehicle counts as they are maintained periodically for enhanced accuracy

Capacity (Traffic) • Maximum theoretical vehicle flow volume

Correlation (Analysis) • A measure of linear association between two variables (Pearson Correlation)

Coverage Area (Traffic) • Physical coverage of the analysis (e.g., county, state)

D

Data Granularity (Data) • The detail in either time or space from a data feed. For example, time granularity may be measured in data records per time interval (e.g., a GPS waypoint every 10 seconds).

Delay (Traffic) • Additional travel time above and beyond the anticipated or expected travel time.

Density (Traffic) • Number of vehicles per mile

E

EMFR (Acronym) • Error to Maximum Flow Ratio

Ensemble Learning (Analysis) • Using multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone

Error to Maximum Flow Ratio (Analysis) • The average of (Absolute value(estimated volume - measured volume)/maximum measure volume). In other terms, the average absolute value of the residual divided by the maximum observed volume on the roadway.

Error to Theoretical Capacity Ratio (Analysis) • The average of

(Absolute value(estimated volume - measured volume)/roadway capacity). In other terms, the average absolute value of the residual divided by the theoretical maximum volume of the roadway.

ETCR (Acronym) • Error to Theoretical Capacity Ratio

F

Factor Groups (Data) • Roads grouped by the type of traffic patterns observed. For example radial freeways differ in traffic patterns to that of beltways.

FHWA (Acronym) • Federal Highway Administration

G

GBM (Acronym) • Gradient Boosting Machine, a machine learning model

GPS Trip / Trajectory / Trace (Data) • A vehicle path defined with an origin/destination and a number of intermediate waypoints, with corresponding time stamps

H

Headway (Data) • Time interval between two consecutive vehicles

HPMS (Acronym) • Highway Performance Monitoring System

Hyper-parameters (Analysis) • Model parameters whose values are set prior to the commencement of the learning process

L

LRS (Acronym) • Linear Referencing System

LSTM model (Acronym) • Long short-term memory neural network – a machine learning model

M

MADW (Acronym) • Monthly volume for Average Day of Week

Map Matching (Analysis) • Reconstructing (most likely) road-based trajectories based on sporadic GPS waypoints

MAPE (Acronym) • Mean Absolute Percentage Error

Mean Absolute Percentage Error (Analysis) • The average of (Absolute value(estimated volume - measured volume)/measured volume). Or expressed in other term, the average absolute value of the residuals divided by the measured volumes.

ML (Acronym) • Machine learning. The field of computer science which gives computers the ability to learn without being explicitly programmed.

MMUT (Acronym) • Mobility Measurement in Urban Transportation

Model Calibration (Analysis) • Using existing volume and other data to come up with a method or algorithm to estimate volumes at places and times for which measured volume is not available.

Model Optimization (Analysis) • Optimize model performance by tuning hyper-parameters

Model Validation (Analysis) • Comparing model estimates with actual observations to ensure its validity

Modeling (Analysis) • Mathematically describing a process or relations.

N

NPMRDS (Acronym) • National Performance Management Research Data Set

O

O-D (Acronym) • Origin-Destination

P

Predictive Analytics (Analysis) • Application of Machine Learning aimed at predicting future events (e.g., future traffic volumes).

Probe Data (Data) • Traffic data derived from probe vehicles equipped with GPS or other technologies

Probe Speed Data (Data) • Speed estimated via vehicle probes

R

R^2 (Analysis) • “R squared” (also called “Coefficient of determination”) is the proportion of the variance in the dependent variable that is predictable from the independent variable(s). It is used to judge the performance of the volume estimator. An R^2 close to 1.0 implies good performance.

Random Forest (Analysis) • A machine learning model

Real-Time Data (Data) • Data feeds with short latency (e.g., 3 min delay), to be used in operations and other immediate point-in-time applications

Recurrent Neural Network (Analysis) • A class of artificial neural networks where connections between units form a directed cycle, which allows it to exhibit dynamic temporal behavior.

Regression Analysis (Analysis) • A set of statistical processes for estimating the relationships among variables. It is often used to predict future events.

Regression Tree (Analysis) • A machine learning method/model based on a tree structure. A decision tree (as a predictive model) to go from observations about an item (represented in the branches) to conclusions about the item’s target value (represented in the leaves).

Residual (Analysis) • For locations where volume data is known,

but that data is not used in calibration, the difference between the estimated volume from the model and the measured volume is the residual.

Road Category (Data) • Type of the road (Interstate, US Route, State route, local road, arterial, collector, etc.)

S

Sample / Capture / Penetration Rate (Analysis) • The percentage of vehicles in a traffic stream that report their speed and location as a probe. Ex. A penetration rate of 10% implies that location and speed data from one out of ten vehicles are processed by a probe data vendor.

Station Data (Data) • Traffic data obtained from a point-based measurement device along a road way, typically some type of sensor to detect a vehicle and measure its speed. This is in contrast to probe data which is derived from vehicle reporting their location and speed.

SVM (Acronym) • Support Vector Machine – a machine learning model

T

Temporal Aggregation (Analysis) • Averages over a specified time interval (eg. 5min, 15min, hour, day)

Test Bed (Analysis) • Within the context of this project, a testbed is volume data from known locations and times from which to both calibrate a volume estimation model, as well as test the model for accuracy.

TMAS (Acronym) • Travel Monitoring Analysis System

TMG (Acronym) • Traffic Monitoring Guide

TTI (Acronym) • Texas Transportation Institute

TTI method (Acronym) • Texas Transportation Institute method for deriving hourly volumes profiles based on AADT data

Turning Movement (Traffic) • Vehicle distribution through interchanges and intersections. Specifically the traffic volume of vehicles that turn left, right or proceed straight on each approach of an intersection.

Vehicle Weight Classes in INRIX trajectory data (Data) • C1<14,000 lbs, C2>14,000 lbs & <26,000 lbs, C3>26,000 lbs

V

Volume (Traffic) • Number of vehicles that pass a particular point (land-

mark) in a given time

Volume Thresholds (Traffic) • Values used to group results

VPP (Acronym) • Vehicle Probe Project

W

Waypoint (Data) • Intermediate location of a vehicle

X

XGBoost (Acronym) • Extreme Gradient Boosting, a machine learning model