For further information, please refer to the TriMet IMI Final Report, Subtask 2.2 Improve Arrival Predictions with Machine Learning Model.

Improve Arrival Predictions with Machine Learning Model

Existing methods for arrival time prediction based on Automated Vehicle Location (AVL) data do not account for changing weather and traffic conditions. Machine learning-based models can improve fixed-route predictions by including external variables and model training.

In 2020, TransLink, the regional transit agency in Vancouver, BC, designed a machine learning-based approach to improve its transit vehicle predictions. TriMet adapted its existing model to produce a similar machine learning model based on TranLink’s process and research. The following focus areas outline the process used to develop a machine learning-based prediction engine:

- Feature Selection and Engineering
- Adapting TransLink’s Model to TriMet’s needs
- Initial Model Training
- Refining Model by Improving Data Inputs
- Pre-Production Machine Learning-Based Model Development
- Comparison of Machine Learning-Based Model to Existing Model
- Model Refinements Based on Analysis Results

The trained TrMet ML models did not produce significantly improved vehicle predictions over non-predicted systems and the system is not currently ready to be put into production use or integrated into TriMet's APIs. The use of other data and/or other statistical models that may perform better are options that may be explored in the future. Some of the work from this effort may be used in these future efforts. Specifically, the development of the ancillary prediction engine components surrounding the core ML models could be adapted for future use.

TriMet’s demonstration included the following feature selections:

Schedule Features	**Weather Features**	**Previous Passage Features**	**Speed/Traffic**
• Month • Day of week • Hour • Minute • Holiday • Scheduled Arrival • Scheduled Time Between Stops • Schedule Deviation	• Rainfall Amount • Snowfall Amount • Temperature • Dew Point • Humidity • Pressure • Wind Speed • Cloud Cover • UV index • Visibility • Drizzle • Presence of Rain • Presence of Snow • Clear	• Time it took last vehicle to transition between stops (Previous Transition) • How long ago previous vehicle was observed to transition between stops (Previous Transition Age)	• Segment has speed data or not • Current speed • Travel time (min) • Speed Age

For further information, please refer to the TriMet IMI Final Report, Subtask 2.2 Improve Arrival Predictions with Machine Learning Model.

The source code Open Transit Tools, the machine learning model is available here:
https://github.com/OpenTransitTools/transitcast
https://github.com/OpenTransitTools/pytransitcast

Data Sources

Schedule Features

Segment Speed (Example)

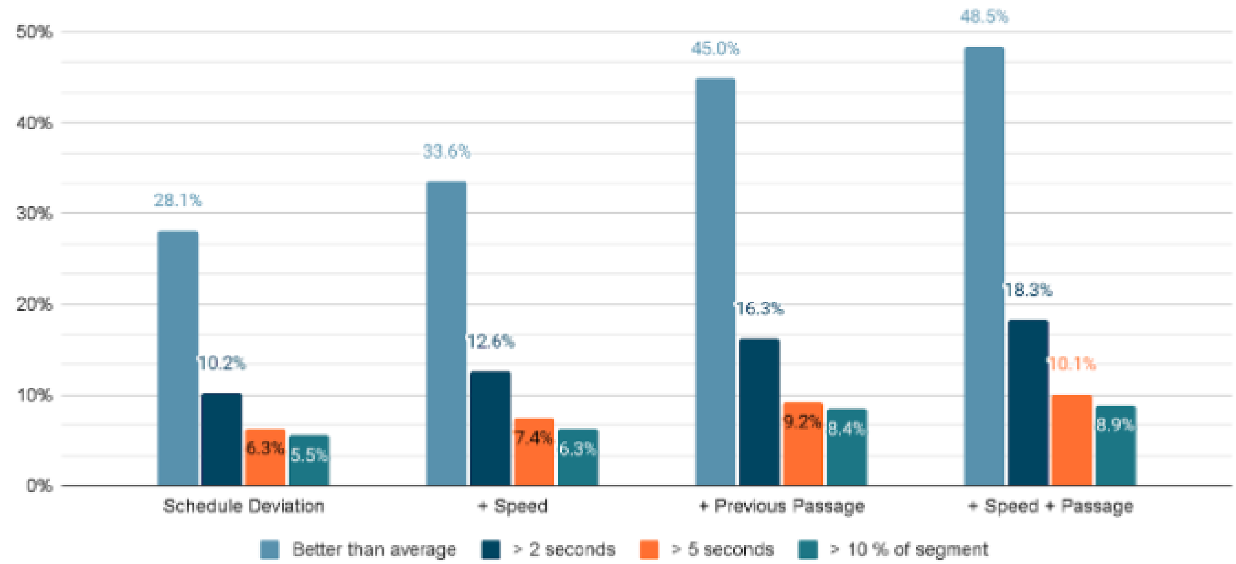
begin_stop_id	end_stop_id	timestamp	has_speed_data	average_speed	confidence	current_speed	free_flow_speed	is_real_time	max_speed	min_speed	number_of_lanes	travel_time_in_minutes	saved_at	updated_at
10000	2672	2022-01-01 08:00:41 UTC	FALSE										2022-01-01 0	2022-01-01 08
10000	9954	2022-01-01 08:00:41 UTC	FALSE										2022-01-01 0	2022-01-01 08
10001	13115	2022-01-01 08:00:02 UTC	TRUE	38	80	44.14	38	TRUE	45	43	2.18	0.2504	2022-01-01 0	2022-01-01 08
10004	10063	2022-01-01 08:00:02 UTC	TRUE	32.3	75.09	34.21	32.3	TRUE	46	24	3.14	0.1519	2022-01-01 0	2022-01-01 08
10005	10004	2022-01-01 08:00:41 UTC	FALSE										2022-01-01 0	2022-01-01 08
10006	7154	2022-01-01 08:00:02 UTC	TRUE	27	99	24	27	TRUE	29	16	1.69	0.4816	2022-01-01 0	2022-01-01 08
10008	10032	2022-01-01 08:00:02 UTC	TRUE	26.62	69.86	24.2	26.62	TRUE	34	17	1.6	0.2817	2022-01-01 0	2022-01-01 08
10008	9014	2022-01-01 08:00:02 UTC	TRUE	29.37	95.16	27.53	29.37	TRUE	34	17	2.55	0.2586	2022-01-01 0	2022-01-01 08
10009	10010	2022-01-01 08:00:02 UTC	TRUE	33	100	32	33	TRUE	38	25	2.17	0.2696	2022-01-01 0	2022-01-01 08
10010	10011	2022-01-01 08:00:02 UTC	TRUE	33	100	32	33	TRUE	38	25	2.17	0.3469	2022-01-01 0	2022-01-01 08

Trip Deviation (Example)

id	created_at	trip_progress	data_set_id	trip_id	vehicle_id	at_stop	delay	deviation_timestamp
157384359	2022-02-01 08:00:03 UTC	59590.31586	10	11282784	3748	FALSE	74	2022-02-01 08:00:01 UTC
157384360	2022-02-01 08:00:03 UTC	-23828.08414	10	11282609	3748	FALSE	74	2022-02-01 08:00:01 UTC
157384362	2022-02-01 08:00:03 UTC	4667.3	10	11282607	4059	TRUE	1	2022-02-01 08:00:01 UTC
157384364	2022-02-01 08:00:03 UTC	-58970.6	10	11282608	4059	TRUE	1	2022-02-01 08:00:01 UTC
157384365	2022-02-01 08:00:03 UTC	82999.3	10	11291829	315	TRUE	80	2022-02-01 08:00:00 UTC
157384366	2022-02-01 08:00:03 UTC	-88790.1	10	11291497	315	TRUE	80	2022-02-01 08:00:00 UTC
157384368	2022-02-01 08:00:03 UTC	0	10	11291833	318	TRUE	0	2022-02-01 07:59:59 UTC
157384369	2022-02-01 08:00:06 UTC	29828.3	10	11292571	306	TRUE	117	2022-02-01 08:00:02 UTC
157384370	2022-02-01 08:00:06 UTC	-3461.3	10	11294632	306	TRUE	117	2022-02-01 08:00:02 UTC
157384373	2022-02-01 08:00:06 UTC	-13845.2	10	11294432	306	TRUE	117	2022-02-01 08:00:02 UTC

INRIX Speed Data

Figure 9. Stop-to-Stop Model Testing Results after Vehicle Position Patch and INRIX Speed Data



Waze Data

Used to explore weather information and correlate this with the mobility trends from other data sources. This data provided interesting insights but wasn't ultimately used in the final MPMs.

GTFS, GTFS-RT and Trip Update Table

https://developer.trimet.org/GTFS.shtml