Our proposed preprocessing algorithm captures temporal features from historical time-series data and extracts spatial characteristics in two scales. The Chicago bike-sharing service operator provides three different datasets of user trips, historical station logs, and bike routes. The user trip dataset indicates bike transitions between two stations in tuples of trip id, timestamp, source station, and destinate station. The algorithm creates a square matrix to efficiently represent the local characteristics of two regions (i.e., stations). The matrix has the size of stations and takes at the row of a source station and the column of a destination. The bike route dataset is a set of GPS coordinates that has a global view of the entire service region by referring to user trip paths. The processing of the bike routes converts the coordinates into non-weighted graphs by taking them as a vertex where edges are generated between regionally proximate. The historical station log dataset contains available bikes and total docks of stations, which are collected every 10 mins. The proposed algorithm pads a missing log with the previous log to preserve temporal features in the station dataset. The processed datasets are separated with window shifting to fit in the input and the output shapes of a model.

An event in our work refers to a state that requires bike rebalancing due to the over-demand of users or over-supply in bike-sharing stations. Our proposed model predicts the events at time and classifies states of the stations into over-demands, normal, or over-supply. The model includes three layers of input, prediction, and output layers. The input layer processes week of the station and trip data to supply daily spatial and temporal features to the prediction layer. The prediction layer computes inputs sequentially by using 3 stacked LSTMs, and the output layer generates prediction results from the last output of the prediction layer with processed bike route data and time . The input layer of the model utilizes convolutional neural networks to capture spatial features from the transition matrix and fully connected neural networks to extract temporal features from the historical station logs. These features in each time-step are converted into a tensor-vector through 2 fully connected neural networks. The LSTMs in the prediction layer has cells to compute the input vectors into a 512-dimensional vector. The output layer takes the one-hot vector of time through a fully connected layer and concatenates it with the 512-dimensional vector. The dimension of the concatenated vector is reshaped into the number of stations through 2 fully connected neural networks. The function activates the reshaped vectors to de-normalize prediction results into the number of bikes in the stations. The static condition method categorizes the results of the output layer into over-demands, normal, or over-supply, where the under and upper boundaries are set as 5% and 95%, respectively.