

# Problem Definition and Characterization

Problem: “Bike Share Demand Prediction”

## Problem Characterization

Stations where bikes are docked and taken from are a common way of implementing a bike sharing system. However, the demand varies greatly between location, weekday and other factors, which leads to imbalances and congestions in the system. This causes further customer dissatisfaction and unreliability of the system. To properly implement dynamic solutions, such as adaptive dynamic pricing, the demand needs to be reliably predicted.

We formulate this as a machine learning problem.

## Modeling Bike Sharing Dynamics

We first assume that customer demand functions are given. We formulate the system dynamics as follows:

- $b_{s(t)}$  : the number of bikes at station  $s \in \mathbb{N}$  at time  $t$

The temporal dynamic of bikesharing is ultimately determined by the demand, which we model using two variables:

- $\text{OutDem}_s(t)$  : the number of bikes that people would take out the dock at  $(t, s)$  if there are enough bikes present
- $\text{InDem}_s(t)$  : the number of bikes that people would dock at  $(t, s)$  if there is be enough space present.

Given demand functions  $\text{InDem}$ ,  $\text{OutDem}$  we model the dynamics using the following assumptions:

- When the out-demand is sufficiently large and there is a sufficient number of bikes, a customer takes a bike out of the station and the number of bikes in the station decreases
- When the in-demand is sufficiently large and there is sufficient space in a station, the customer docks a bike in the station

This defines the temporal evolution of the number of bikes in the system, starting at  $B(t = 0)$ :

$$B_s(t + 1) = B_s(t) + \min(\text{InDem}_s(t), \text{Cap}_s(t)) - \min(\text{OutDem}_s(t), B_s(t))$$

## Possible Modelling Approaches

### Problem Characteristics

Both quantities  $\text{InDem}$  and  $\text{OutDem}$  appear implicitly in the model. These quantities model the combination of psychological and logistical demand of the population to use a bike for travelling from a start station to a target station at a given time. From common sense reasoning, we expect the motivation to use a bike on:

- the current day in the week,
- the time in the year, determining factors such as expected temperature and general conditions,
- the daytime, determining e.g. direction of workers travelling from and to their workplaces,
- the weather

Fundamentally there are also short-term temporal dependencies, i.e. given that many people have used a shared bicycle to reach their workplace at a specific day, they are more likely to also return to their home. Another dependency arising is the connectivity of stations, i.e. if the demand to take a bike rises at one station, the demand to dock a bike at a **connected** station (i.e. there occur frequent rides between these two stations) should rise.