Portfolio scripts

**1 AIRMAP**

We obtained a functional relationship between atmospheric PM2.5 and land use pattern based on multiple linear regression model. In this model, we extract features from various geographic information data (population density, road length, wind speed, etc. ). We applied the model to Beijing, Tianjin, Hebei provinces in China and visualize the results with interactive map.

**Motivation**

Poor air quality is one of the greatest environmental hazards facing many urban areas around the globe. PM2.5 is the designation for particulates that are less than 2.5 um in aerodynamic diameter, making it small enough to be inhaled.

The concentration of PM2.5 depend on location, but there is a larger variability in and around highly developed urban areas[1]. Currently, we obtain air quality data from discrete air quality monitoring stations. To get continuous aerosol concentration, we extract features for spatial interpolation using various geographic information data.

**Regression model construction with Python**

**Data visualization with echarts.js**

**UI design with Photoshop**

**Map1**

**Data of discrete air quality monitoring stations**

**Map2**

**Interactive continuous air quality map**

**Analysis Mode**

**Compare daily air quality in month’s scale**

**2 Bacterial Communities**

**3 Melody Flow**

This project is based on ‘Magenta’ , an open-sourced tool for machine learning in the creative process, developed by Google Brain. We tend to use RNN to generate music in different genres.

In this project, we trained models with light music, popular music and classical music. The model we used including (1) traditional LSTM; (2) LSTM with LookBack mechanism; (3) LSTM with Attention mechanism.

(You can find pieces of music we generated in the result section.

**Motivation**

Inspired by the idea of style transfer in images, we hope to extract features in different genres of music and generate music by using several given notes as primer. We choose Recurrent Neural Network (RNN) as the basis of our model, as is suitable for extracting features in time sequence.

**Work Flow**

The right side chart shows our basic work flow.

We use every note as an input vector. The note is featured by its start time(Note-on), end time(Note-off), The next occurrence time(Time-shift),pitch and velocity and constructed into a 1416-dimensional vector after onehot encoding

**Results**