

机器学习回顾 & Location 实验

游浩然

u201515429@hust.edu.cn

June 19, 2018

Table of Contents

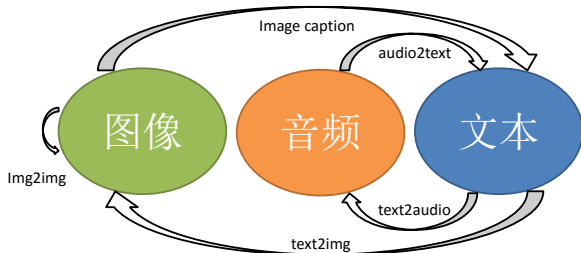
I. 机器学习回顾

- 优化问题
- 分/聚类问题
- 预测问题

II. Location Assignment

机器学习

- 传统机器学习
 - 优化问题
 - 分/聚类问题
 - 预测问题
- 深度学习



优化问题

Example: EV Charger Location for Ireland (ICM 2018)

- destination charging designed for charging for several hours at a time or even overnight.
- supercharging designed for longer road trips to provide up to 170 miles of range in as little as 30 minutes of charging.

one of five tasks:

- Determine the optimal number, placement and distribution of charging stations if your country could migrate all their personal passengers vehicles to all-electric vehicles instantaneously.
- Present a proposal for evolving the charging network of your chosen country from zero chargers to a full electric-vehicle system.
- Based on your growth plan, what is the timeline you propose for the full evolution to electric vehicles in your chosen country.

分/聚类问题

Example: 对通讯记录进行社群划分 (华中赛 2017)

- 给定无向图，区分个体差异，实现分群
- 给定有向图，刻画任意两点间的相似关系
- 给出某营业部近三个月的内部通讯记录，结合并完善问题 1, 2 的数学模型，实现个体的分群。
- 考虑通讯的位置，时间，通话频率等多种因素建立综合的数学模型，挖掘更多的信息。

预测问题

Example: 高炉炼铁-铁水含硅量预测 (MathorCup 2017)

炼铁过程依时间序列采集的工艺参数是一个高维的时间序列，影响因素数以百计。其终极生产指标产量、能耗、铁水质量等指标都与炼铁过程中的一项控制性中间指标——炉温，即铁水含硅量 $[\text{Si}]$ 密切相关。

- 从给定数据表中 $[\text{Si}]$ - $[\text{S}]$ -FL-PML 依序号排列的 1000 炉生产数据中，自主选择学习样本和算法，建立 $[\text{Si}]$ 预测动态数学模型，包括一步预测和两步预测。
- 自主选择验证样本，验证预测成功率，包括数值预测成功率和炉温升降方向预测成功率，并讨论其动态预测控制的可行性。
- 建立质量指标 $[\text{S}]$ 的优化数学模型，讨论按照优化模型计算结果进行 $[\text{Si}]$ 预测控制的预期效果。

Table of Contents

I. 机器学习回顾

II. Location Assignment

- Method
- Result

KNN

main-idea:

- 用 K 个最近邻样本输出的平均值作为回归预测值。

feature selection:

- 取训练集和测试集 feature 的交集，总特征数为 423
- 将信号强度 0 变为极弱信号强度-100
- 按比例划分训练集和交叉验证集。

Distribution

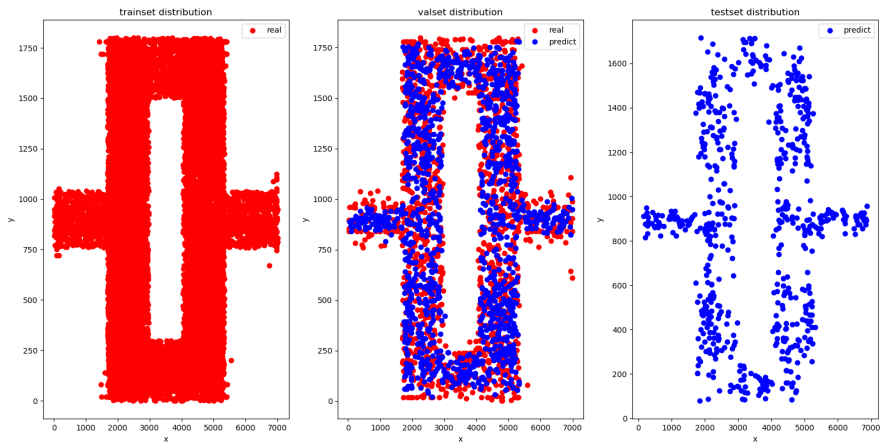


Figure 1: Distribution

CDF

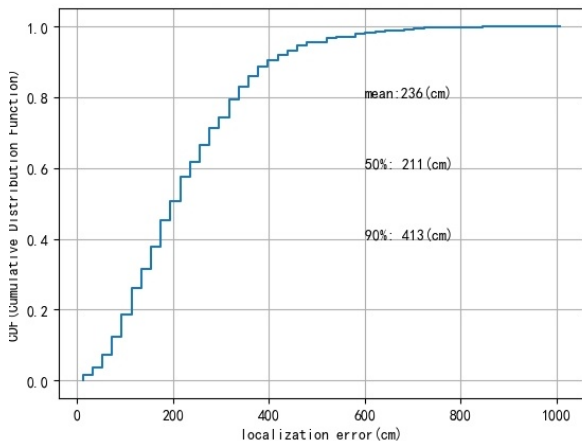


Figure 2: Cumulative Distribution Function

Thank You.