

智慧城市交通大数据分析模型研究

摘要

本文研究背景是基于智慧城市交通的大数据分析，主要要实现基于智慧城市交通数据的大数据分析以及建立基于智慧城市交通数据的大数据分析模型，包含以下内容：

- 1、对大数据分析工具的熟练使用。能对数据进行分析、数据建模等等
- 2、应用大数据分析技术，将长期积累下来的各种技术指标数据形成可视化，以此来识别城市实时的交通条件、评估点到点的耗费时间、提供方便快捷路径的建议等，这些将极大的改善交通拥堵状况，提高道路的的畅通性安全性，减少事故的发生，提供便携出行。
- 3、建立合适的模型以深入分析和挖掘，发挥数据的协同作用。

目前，对于研究方法的初步设想是运用大数据分析工具将与智能交通相关的多源数据分析，数据建模，深入分析，以改善交通拥堵等交通问题。

本文的主要内容如下：

我首先详细地说明了智慧城市交通在国内外的研究现状和发展前景，从我国交通的发展现状、城市化与信息化对交通的影响开始阐述，具体说明了我国交通发展中遇到的挑战，随后阐述了交通发展观念的变化发展，最后点明智慧交通的国内外发展现状。

其次我又说明了本次设计过程中将会用到的大数据分析相关技术，查阅了许多相关资料，了解到 python 中有 numpy,pandas,scipy, sklearn 等库,非常适用于做数据分析，适用于本次设计，所以我最终选用 python 语言，作为本次设计的编程语言和环境。

然后我开始研究分析从老师那里获取的交通数据集，基于种种原因，我最终选用高架指数，地面指数和交通事故数据这几个数据集，分别进行探索分析并进行可视化，主要分析地面交通拥堵情况、高架拥堵情况、交通事故类型和经常发生事故的地段，并通过可视化结果，详细地分析交通拥堵和时间的关系、事故类型分布等。

最后，我将这些数据进行预处理后建模，并评估模型。通过对数据分析过程的研究，最终使用回归模型对本次设计中的高架指数、地面指数数据集进行建模。我以指数为因变量，其他数据列——主要是时间作为自变量，建立指数预测模型。并进行评估，选出相对最优的模型。最终发现高架指数数据集与梯度提升回归树模型拟合度最高，梯度提升回归树是高架指数的最优模型。而地面指数数据集与随机森林模型拟合度最高，随机森林模型是地面指数的最优模型。

关键字：智慧交通；数据分析；数据建模

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。

如要下载或阅读全文，请访问：

<https://d.book118.com/495242014104012010>