### **项目名称：基于自然语言处理的城市交通舆情意图识别研究与实现**

#### **一、 项目简介 (Introduction)**

本项目旨在应用前沿的自然语言处理（NLP）技术，解决城市交通领域中公众意见繁杂、难以快速归纳分析的痛点。社交媒体平台每日都会产生大量关于交通路况、公共交通服务的文本数据，这些数据蕴含着宝贵的民生信息。本项目通过构建一个智能分类系统，能够自动读取并理解这些文本，将其精准地识别为预定义的意图类别（如“路况投诉”、“服务建议”等），从而为交通管理部门提供高效、精准的数据洞察，辅助其进行科学决策。

#### **二、 核心流程 (Core Workflow)**

项目遵循标准的NLP任务解决流程，主要分为以下五个阶段：

1. **数据准备 (Data Preparation):**
   * 从社交媒体等公开渠道收集与城市交通相关的原始文本数据。
   * 对数据进行清洗，去除无关信息，并在此基础上进行人工标注，构建一个高质量、带意图标签的领域数据集。
2. **模型选择 (Model Selection):**
   * 选用一个强大的中文预训练语言模型（如BERT或MacBERT）作为技术基座。这类模型已在海量文本上进行过学习，具备了深厚的中文语法和语义理解能力。
3. **模型微调 (Model Fine-tuning):**
   * 将准备好的标注数据集“喂”给预训练模型，进行针对性的“二次训练”。在这个过程中，通用模型将学习并适应交通领域的特定语言模式和分类任务，成为一个交通意图识别的“专家模型”。
4. **效果评估 (Performance Evaluation):**
   * 在独立的测试数据集上，对微调后模型的分类准确性、可靠性进行全面评估，确保其达到预期的应用标准。
5. **应用展示 (Application Demonstration):**
   * 开发一个简单的交互式Web界面，用户可以输入任意交通相关的句子，系统能够实时返回其预测的意图类别，直观地展示项目的最终成果。

#### **三、 实现方法简述 (Methodology)**

* **数据处理:** 使用 pandas 库对数据进行整理和分析。意图标签体系经过精心设计，以确保其既能覆盖大部分场景，又相互独立。
* **模型实现:** 整体技术栈基于 PyTorch 深度学习框架。通过调用 Hugging Face Transformers 库，可以方便地加载预训练模型并利用其高级API（如 Trainer）来执行微调和评估流程。
* **前端演示:** 利用 Streamlit 库快速搭建Web应用，它能用纯Python代码构建出简洁美观的交互界面，非常适合AI项目的原型展示。