**需求分析报告**

**项目概述**

本报告针对的是一款基于甲言开源大模型的Web端文言文处理工具。该工具旨在为用户提供便捷的文言文分词、断句、词性标注以及标点功能。通过结合现代自然语言处理技术与古代语言特点，该工具将帮助用户深入理解和分析文言文。

**功能需求**

**1. 文言文分词**

* **需求描述**：用户输入一段文言文文本，系统应能够对文本进行准确分词。
* **实现方式**：通过后端服务器上的CharHMMTokenizer模型实现。
* **用户界面**：提供输入框供用户输入文言文文本，以及一个“分词”按钮触发分词操作。

**2. 词性标注**

* **需求描述**：对用户输入的文言文文本进行词性标注，帮助用户理解每个词的语法功能。
* **实现方式**：利用CRFPOSTagger模型进行词性标注。
* **用户界面**：提供“词性标注”按钮，用户点击后，系统展示标注结果。

**3. 文言文断句**

* **需求描述**：自动识别文言文文本中的句子边界，进行断句处理。
* **实现方式**：通过CRFSentencizer模型完成断句。
* **用户界面**：提供“断句”按钮，用户点击后，系统展示断句结果。

**4. 文言文标点**

* **需求描述**：在用户输入的文言文文本中添加适当的标点符号。
* **实现方式**：利用CRFPunctuator模型进行标点预测。
* **用户界面**：提供“标点”按钮，用户点击后，系统展示添加标点后的文本。

**技术需求**

**后端**

* 使用Python编程语言和Flask框架构建Web服务。
* 集成Jiayan开源自然语言处理模型以支持文言文处理功能。
* 定义RESTful API端点以接收前端请求并返回处理结果。

**前端**

* 使用HTML、CSS和JavaScript构建用户界面。
* 前端应具备响应式设计，以支持不同设备和屏幕尺寸。
* 通过AJAX技术与后端进行异步数据交互。

**部署**

* 确保Web服务可以部署在标准的服务器环境中，例如Apache或Nginx。
* 应支持HTTPS协议，保障数据传输安全。

**用户界面需求**

* 提供一个简洁、直观的界面，允许用户输入文言文文本。
* 显示处理结果，并允许用户对比原始文本和处理后的文本。
* 提供错误处理机制，例如在用户未输入文本或服务器处理错误时给出明确的提示。

**性能需求**

* 确保所有处理功能响应时间在可接受范围内，通常不超过数秒。
* 系统应能够处理不同长度的文言文文本，而不会出现明显的性能下降。

**安全需求**

* 对用户输入进行验证和清理，防止注入攻击。
* 确保后端服务运行稳定，并进行适当的异常处理。

**维护和扩展性需求**

* 代码应具有良好的结构，以便于后续维护和功能扩展。
* 提供文档说明系统架构、安装部署过程以及API使用方法。

**结论**

本文言文处理工具将极大便利学术界和普通用户对文言文资料的研究与学习。通过实现上述功能需求，项目将填补市场对高效、准确的文言文在线处理工具的空缺。