Qualcomm Developer Project

LLaMA on WoS

Project Submission

|  |  |  |
| --- | --- | --- |
| **Attribution:** | Tianye Fan | |
| **Email address** | <zhangzz6687@thundersoft.com>  [tianye.fan@thundersoft.com](mailto:tianye.fan@thundersoft.com) | |
| **Project Title**\* | LLaMA on WoS | |
| **Images**  *Upload up to 5 images of your project*  *Please submit/send the original JPEG/PNG files for all images included in the document* | **WoS**    **Type-c usb line**  **typc**  **Charger**  charger | |
| **Description**\*  *High level description of the project* ***(75 words or less)*** | The project builds a CLI application that runs the open-source model LLaMA v2-7B on the WoS platform, enabling interactive multi-turn dialogue inference based on QNN. | |
| **Objective**   * *What inspired you to create this project?* * *What is your desired outcome?* | LLaMA is an open-source conversational reasoning model that provides intelligent answers and suggestions through interactive multi-turn dialogues with users. This project aims to port the LLaMA model to the WoS platform, exploring its potential for practical applications on mobile devices and providing users with a more convenient and efficient conversational experience. | |
| **Materials Required / Parts List / Tools** | Part Name | Link to purchase |
| Wos | https://www.qualcomm.com/support/contact/forms/snapdragon-developer-kit |
| USB Line | https://item.jd.com/40759941966.html |
| Charger | https://www.qualcomm.com/support/contact/forms/snapdragon-developer-kit |
| **Source Code / Source Examples / Application Executable**  *Link to open source / shareable code repository* | Description | Link |
| Source Code | https://github.com/ThunderSoft-XA/QCS8380-LLaMA-on-WoS |
| **Additional Resources**  *List related links or resources such as websites, videos, presentations, or other materials* | Resource Title | Link or File Name (and provide file) |
| Video | https://github.com/ThunderSoft-XA/ QCS8380-LLaMA-on-WoS/doc/usage.mp4 |
| **Build / Assembly Instructions** | The project is released in binary form and no need to build. For Qualla build documentation, please refer to <https://github.com/ThunderSoft-XA/QCS8380-LLaMA-on-WoS/doc/build.md>. | |
| **Usage Instructions** | Refer to [https://github.com/ThunderSoft-XA/ QCS 8380-LLaMA-on-WoS/doc/user-guide.md](https://github.com/ThunderSoft-XA/C8380-LLaMA-on-WoS/doc/user-guide.md). | |
| **Contributor(s) Info**  *Feel free to include headshots!* | Name | Title  Company |
| <zhangzz6687@thundersoft.com> | Thundersoft |
| [tianye.fan@thundersoft.com](mailto:tianye.fan@thundersoft.com) | Thundersoft |

––– Continued on next page –––

Filters and Tags for QDN projects page

|  |  |  |
| --- | --- | --- |
| **Platform/Hardware** | CSR 101x/102x Bluetooth  DragonBoard 410c  mangOH Red/Yellow  Qualcomm C6490P | MDM920x LTE for IoT  QCA-402x WiFi/BLE/Zigbee  Qualcomm Robotics RBx Dev Kit  √ Qualcomm QCS8380 |
| **Software Tools** | 3D Audio Plugin for Unity  Adreno GPU SDK  Hexagon DSP SDK | Neural Processing SDK for AI  √ 　Qualcomm AI Engine Direct  　Snapdragon Profiler |
| **Operating System** | Android  Linux  ThreadX RTOS | Ubuntu Core  √ Windows 10 IoT Core |
| **Cloud Services/Platform** | Sierra Wireless AirVantage  Gizwits Cloud Platform  AT&T M2X  IBM Bluemix | IBM Watson IoT  Microsoft Azure IoT  Amazon AWS IoT |
| **Skill Level Required** | Advanced  Beginner  √ Intermediate |  |
| **Areas of Focus** | 3D Printing & Modeling  Alexa Voice Service  √ Artificial Intelligence  Bluetooth  √ Computer Vision  Digital Signage  Education  √ Embedded  Gaming | Healthcare  IoT  Robotics  Security  Sensors  Smart Cities  Smart Home  Toys |

*By submitting your content (“Submission”), you are granting Qualcomm a royalty-free, perpetual, non-exclusive, unrestricted, worldwide license to: (a) post, use, copy, sublicense, adapt, transmit, publicly perform or display any such Submission, (b) use, reproduce, modify, adapt, publish, translate, create derivative works from, distribute, perform, play, host, communicate, make available and publish your Submission without restriction and (c) sublicense to third parties the unrestricted right to exercise any of the foregoing rights granted with respect to the Submission. The foregoing grants shall include the right to exploit any ideas, concepts, intellectual property, or proprietary rights in such Submission, including but not limited to rights under copyright, trademark, servicemark or patent laws under any relevant jurisdiction without Qualcomm owing any monies to you whatsoever. You represent and warrant that you own all right, title and interest in and to the Submission, or you have been granted sufficient rights in and to the Submission allowing the foregoing use of such Submission.*