Literature Review:

1. **Introduction:**
   Predicting poverty levels using socioeconomic indicators is a significant area of research that addresses key challenges in resource allocation and policymaking. Accurate predictions of poverty can guide targeted interventions, aligning with global goals like Sustainable Development Goal 1: No Poverty. A review of existing literature is essential to understand the methodologies, findings, and gaps in prior studies. This helps contextualize the current project and identify areas where it can contribute to the existing body of knowledge.

2. **Organization:**
   This literature review is organized thematically, grouping studies based on their focus on machine learning techniques, socioeconomic datasets, and poverty prediction methodologies.

3. **Summary and Synthesis:**
   Machine Learning Techniques for Poverty Prediction
   - **H. Zixi (2021)**: In "Poverty Prediction Through Machine Learning," Zixi applied traditional machine learning models, such as Random Forest and Gradient Boosting, to predict poverty using structured data. The study demonstrated the efficiency of machine learning models in identifying key poverty indicators but highlighted challenges in handling missing data and ensuring generalizability across regions.
   - **Aziza Usmanova**: In her review, Usmanova explored various artificial intelligence utilities in poverty prediction. She emphasized the importance of feature selection and preprocessing for improving model accuracy and discussed the potential of hybrid models in combining traditional and deep learning approaches.

   Socioeconomic Datasets in Poverty Research
   - **World Development Indicators (WDI)**: The WDI dataset from the World Bank is widely used in poverty studies due to its comprehensive coverage of indicators like GDP, literacy rates, and healthcare access. Previous research has shown that higher GDP per capita and literacy rates correlate with lower poverty levels. Studies also emphasize the need for robust preprocessing, including imputation of missing data and feature scaling, to ensure model compatibility.
   - **Comparative Studies**: Research contrasting datasets like WDI with satellite imagery data has shown that combining diverse data sources can enhance prediction accuracy, though it increases computational complexity.

   Key Methodologies and Their Applications
   - **Scikit-learn**: Zixi's study demonstrated the utility of Scikit-learn for traditional machine learning models. Its simplicity and interpretability make it a preferred tool for studies with structured datasets like WDI.
   - **TensorFlow and Keras**: While primarily used in deep learning, these tools have been explored in poverty prediction studies for their ability to handle unstructured data. However, they are resource-intensive and require significant computational power.

4. **Conclusion:**

The literature reviewed highlights the effectiveness of machine learning in predicting poverty, the importance of datasets like WDI, and the need for rigorous preprocessing. However, gaps exist in tailoring models to specific socioeconomic contexts and combining diverse data sources. This project aims to address these gaps by leveraging the WDI dataset with machine learning techniques to develop interpretable and accurate models. By doing so, it will contribute to the body of knowledge by offering a scalable and context-sensitive approach to poverty prediction.

5. **Proper Citations:**

- Zixi, H. (2021). "Poverty Prediction Through Machine Learning." 2021 2nd International Conference on E-Commerce and Internet Technology (ECIT), Hangzhou, China, pp. 314-324. DOI: 10.1109/ECIT52743.2021.00073.
- Usmanova, A. "Utilities of Artificial Intelligence in Poverty Prediction: A Review."
- World Bank Development Indicators (WDI): https://databankfiles.worldbank.org/public/ddpext/.
- Scikit-learn: https://scikit-learn.org/stable/.
- TensorFlow: https://www.tensorflow.org/.