Assignment Guide: Breast Cancer Data Analysis and Streamlit App

Step 1: Project Setup

1. **Create a Project Directory:**
    - Create a new directory for your project in VS Code.
    - Initialize a Git repository.
2. **Set Up a Virtual Environment:**
    - Create and activate a virtual environment for the project.

Step 2: Dataset Acquisition and Preparation

1. **Download the Dataset:**
    - Download the Breast Cancer dataset from a reliable source like the UCI Machine Learning Repository, Kaggle or get the dataset from sklearn.
2. **Data Preparation:**
    - Write a Python script to load and preprocess the dataset, ensuring it is ready for analysis.

Step 3: Feature Selection

1. **Feature Selection Technique:**
    - Implement feature selection using methods like `SelectKBest` from `sklearn.feature_selection`.

Step 4: Grid Search CV for Model Tuning

1. **Grid Search Cross-Validation:**
    - Provide a template or guide for setting up Grid Search CV to optimize the parameters of an ANN model (`MLPClassifier` from `sklearn.neural_network`).

Step 5: Implementing an Artificial Neural Network (ANN) Model

1. **ANN Model Creation:**
    - Outline the steps to create an ANN model.
    - Train and evaluate the model using the breast cancer dataset.

Step 6: Building a Streamlit App Locally

1. **Streamlit code:**
    - Use Streamlit as a tool for building interactive web apps with Python.
2. **Developing the Streamlit App:**
    - Create a basic Streamlit app that allows users to interact with the breast cancer dataset and view model predictions.
    - Integrate model predictions, and user interaction within the Streamlit app.

Step 7: Deployment and Version Control

1. **GitHub Repository Setup:**
    - Setting up a GitHub repository for their project. Give the link in the comment section.
    - Commit their code regularly and push changes to GitHub.
2. **Submission Requirements:**
    - Specify the deliverables, such as the Python scripts, Streamlit app code, and a README.md file documenting the project.

Additional Tips

- **Documentation and Comments:** Emphasize the importance of clear documentation and comments in the code to explain each step and rationale.
- **Encourage Exploration:** Encourage students to explore different feature selection techniques, model architectures, and hyperparameter configurations beyond the basic requirements.

By following these steps, students can gain hands-on experience in data preprocessing, model development, and interactive web application creation using Streamlit, enhancing their understanding of machine learning concepts and practical skills.