JON MARE EDRIC P. LONTAYAO **BSIT4B**

"Choosing the Optimal Linux Distribution for an AI Development Server"

Scenario Description

- The system is a dedicated server for training and deploying machine learning models.
- It needs GPU acceleration, fast package management, and support for AI frameworks such as TensorFlow and PyTorch.
- The server must remain secure and stable for long-term use.

Options Considered

1. **Ubuntu (20.04/22.04 LTS)**
    - Pros: Excellent NVIDIA CUDA and AI framework support, large community, stable LTS updates.
    - Cons: Slightly heavier base installation.
2. **Fedora Workstation / Server**
    - Pros: Cutting-edge software, frequent updates, modern kernel support.
    - Cons: Shorter support cycle (about 13 months).
3. **Debian Stable**
    - Pros: Very stable, low resource usage, excellent security reputation.
    - Cons: Older packages — requires manual backports for latest AI tools.

Evaluation Criteria

- **Performance:**
  Ubuntu and Fedora both ship modern kernels with excellent GPU driver support, while Debian may lag behind.
- **Security:**
  Debian has the strongest reputation for security, but Ubuntu LTS also provides timely patches and long-term security maintenance.

- **Maintenance:**
  Ubuntu LTS offers 5 years of official support, making it easy to maintain. Fedora requires frequent upgrades, while Debian is stable but less suited for rapidly evolving AI libraries.

Selected Distribution

Ubuntu 22.04 LTS (Long-Term Support)

Justification

- Performance: Optimized for AI workloads, seamless NVIDIA CUDA and cuDNN installation, and direct support from major AI frameworks.
- Security: Regular security patches with Canonical's long-term support guarantees.
- Maintenance: 5-year support lifecycle minimizes upgrade frequency, and apt package manager simplifies updates.
- Community & Documentation: Vast online resources and tutorials tailored to AI and ML use cases.

Conclusion

Ubuntu 22.04 LTS is the optimal Linux distribution for an AI development server due to its balance of performance, security, and long-term maintainability. Its strong ecosystem ensures minimal setup effort and maximum reliability for high-compute workloads.