1. The system shall respond quickly under heavy user load.

2. The model shall achieve a minimum recall of 0.85 on the medical test dataset.

3. The application should be intuitive and user-friendly.

4. The monitoring system shall flag drift if PSI exceeds 0.25 for any critical feature.

5. The latency should be minimized whenever possible.

6. The ML pipeline shall log all data preprocessing steps with timestamps and version identifiers.