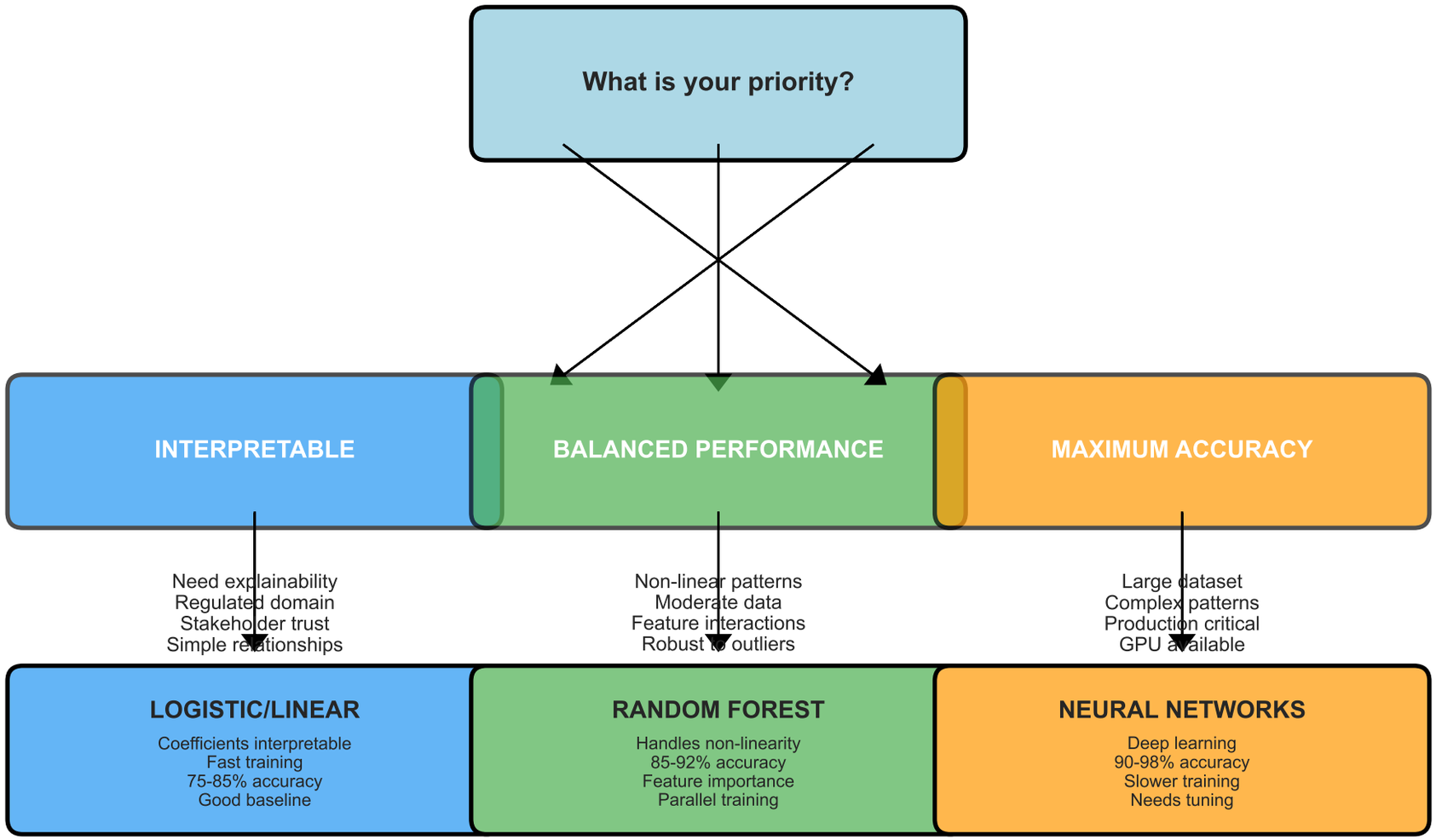
When to Use Which Classification Algorithm: Decision Framework



Additional Considerations

```
Class Balance: Severe imbalance (>95:5) → Use SMOTE/class weights or ensemble methods
Feature Count: <20 features → Logistic sufficient; 20-100 → Random Forest; >100 → Neural nets
Linearity: Linear separable → Logistic/SVM; Complex boundaries → Trees/Neural nets
Training Time: Real-time updates → Online logistic regression; Batch → Any method viable
Deployment: Edge devices → Small models (logistic, small trees); Cloud → Large ensembles OK
Multi-class: 2 classes → All methods; >10 classes → Neural nets or hierarchical classification
```

Principle: Start interpretable (logistic), add complexity (trees/SVM) only when accuracy demands it