Name – A.Shruthi

2303A52133

Batch – 37

RandomForestClassifier is trained on sklearn.datasets.load_digits() features (64 pixels). Permutation Importance measures the drop in accuracy when pixel values are permuted and is visualized as an 8×8 heatmap. SHAP (TreeExplainer) provides both global (beeswarm) and local (force) explanations showing feature contributions per class and per instance. LIME (LimeTabularExplainer) explains individual predictions by approximating the model locally and mapping feature weights back onto the 8×8 pixel grid.

What the notebook produces (deliverables)

- outputs/pi_heatmap.png — Permutation Importance 8×8 heatmap.
- outputs/shap_beeswarm_aggregated.png — SHAP global beeswarm (aggregated).
- outputs/shap_force_digit3.html (or a fallback PNG) — SHAP local force plot for a chosen digit '3'.
- outputs/lime_digit3.png and outputs/lime_digit8.png — LIME explanations mapped to 8×8 for two test instances.
- outputs/rf_digits_model.joblib — saved RandomForest model.
- Clean, runnable Python script / notebook in the canvas (open **Digits Pi Shap Lime Notebook**).

5–10 comparative insights (consistencies & differences)

1. Permutation Importance highlights pixels that, when shuffled, degrade overall model performance — it tends to mark *globally* important regions (center strokes) but doesn't show direction (positive/negative) of influence.
2. SHAP gives signed, additive contributions per class: you can see which pixels push the model *towards* or *away from* class 3 (or any other class) at instance-level and globally (beeswarm). SHAP often aligns with PI for top pixels but adds directionality.
3. LIME explains the prediction locally using linear approximations; its highlighted pixels often match SHAP's local top contributors but can differ because LIME perturbs and fits a local surrogate while SHAP uses game-theoretic attributions.
4. PI is robust for finding generally important pixels across the whole test set; SHAP is better for per-class and per-instance nuance; LIME is useful for quick, interpretable local checks but can be unstable to perturbation settings.
5. For some digits (e.g., '3' vs '8'), SHAP may show opposite-signed contributions in overlapping pixels (the same pixel supports '3' but suppresses '8'), while PI will mark that pixel important for overall predictions but won't tell you the sign.

6. LIME maps often emphasize fewer features (sparser) because you request a limited num_features — good for presentation but might miss subtle contributors SHAP reports.
7. If the RandomForest relies on combinations of pixels (nonlinear interactions), PI may capture importance but not interactions; SHAP can show approximate interaction patterns if you inspect interaction values (optional extension).
8. Practical tip: use PI first to find candidate important regions, then use SHAP for global/class-level interpretation and LIME for human-friendly local explanations when presenting to non-technical audiences.
9. Expect small disagreements in feature ranking between methods — these are diagnostic, not errors: they reflect different definitions of "importance" (performance drop vs. additive contribution vs. local linear fit).