```
$ python aggregate_summaries.py --root artifacts --csv all_results.csv
=== Aggregated Results (sorted by accuracy desc) ===
   ACC      F1  MODEL                RUN
-----------------------------------------
0.8535  0.8534  RandomForest         artifacts\rf_n400_d16_sqrt
0.8495  0.8495  RandomForest         artifacts\rf_n600_d12_sqrt
0.8450  0.8450  RandomForest         artifacts\rf_n200_d12_sqrt
0.8425  0.8425  RandomForest         artifacts\rf_n400_d12_sqrt
0.8380  0.8379  RandomForest         artifacts\rf_n400_d12_log2
0.7695  0.7695  DecisionTree         artifacts\dt_d16
0.7625  0.7624  DecisionTree         artifacts\dt_d10
0.7590  0.7587  DecisionTree         artifacts\dt_d12
0.7570  0.7562  DecisionTree         artifacts\dt_d8
0.7535  0.7526  DecisionTree         artifacts\dt_d6
CSV written to: all_results.csv
```

Best decision tree was d16 with .7695 accuracy. Best random forest was n400 d16 with .8534 accuracy.

Confusion matrices:

Decision Trees

Depth 6:

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	693	307
Actual: Not Bee	186	814

Depth 8:

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	699	301
Actual: Not Bee	185	815

Depth 10:

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	738	262
Actual: Not Bee	213	787

Depth 12:

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	723	277
Actual: Not Bee	205	795

Depth 16:

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	755	245
Actual: Not Bee	216	784

Random Forests:

n200 d12

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	854	146
Actual: Not Bee	164	836

n400 d12

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	861	139
Actual: Not Bee	185	815

n400 d16

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	855	145
Actual: Not Bee	170	830

n400 d12 log2

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	884	116
Actual: Not Bee	177	823

n600 d12

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	862	138
Actual: Not Bee	163	837

Part C

```
Test classification report:
              precision    recall  f1-score   support

       NOBEE       0.96      0.91      0.93      1000
         BEE       0.91      0.96      0.94      1000

    accuracy                           0.93      2000
   macro avg       0.94      0.93      0.93      2000
weighted avg       0.94      0.93      0.93      2000
```

	Pred: Bee	**Pred: Not Bee**
Actual: Bee	910	90
Actual: Not Bee	41	959

Part D

Model	Accuracy	Macro F1
Decision Tree	.7695	.7695
Random Forest	.8535	.8534
Shallow CNN	.94	.93

Random forests improve over decision trees by combining predictions from multiple trees, averaging out noise and stabilizing predictions. Decision trees tend to overfit, learning noise and patterns that don't generalize well.

CNNs improve over RFs because they're better at recognizing shapes. By recognizing local edges and textures, it more accurately identifies the bees. Translation tolerance allows the network to identify the bees even when they've shifted position, and global pooling helps to summarize features, giving it stronger generalization for classification tasks.

In the confusion matrices, both the DTs and RFs make a moderate number of errors, with RFs having fewer. CNNs on the other hand have fewer misses. In terms of the effect false negatives and false positives would have on data, missing bees would distort population monitoring, and false alarms would waste time on non-bee detection.