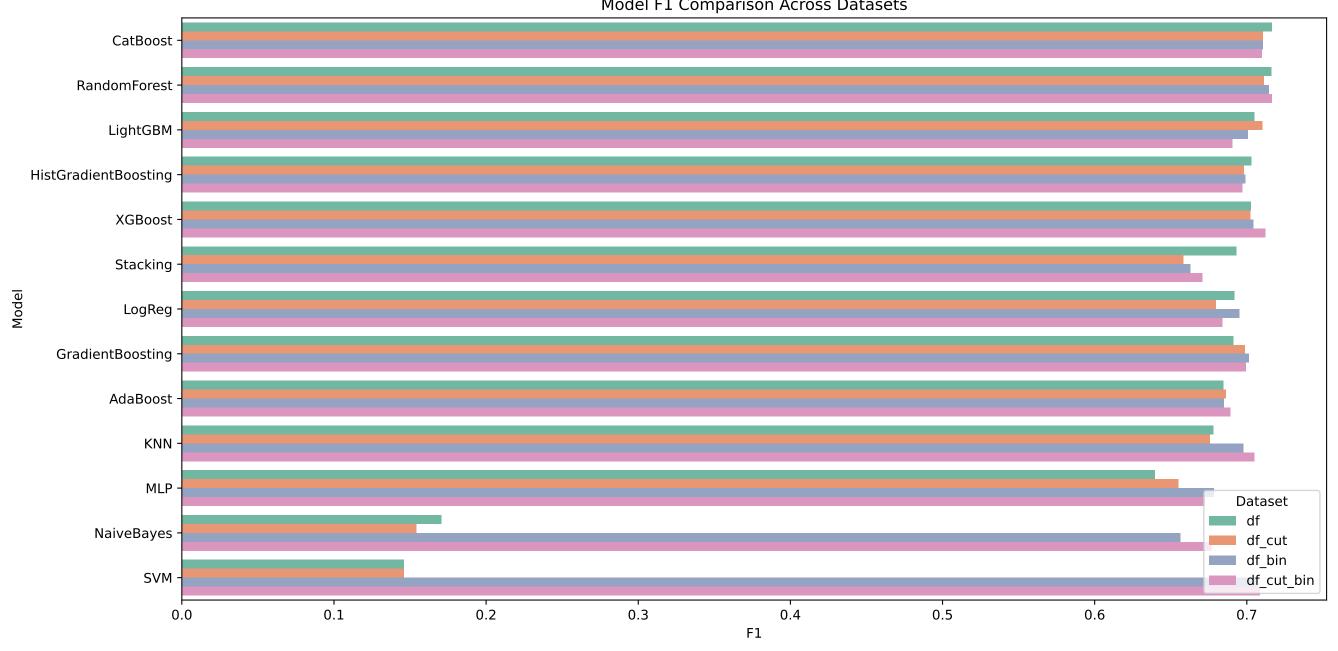
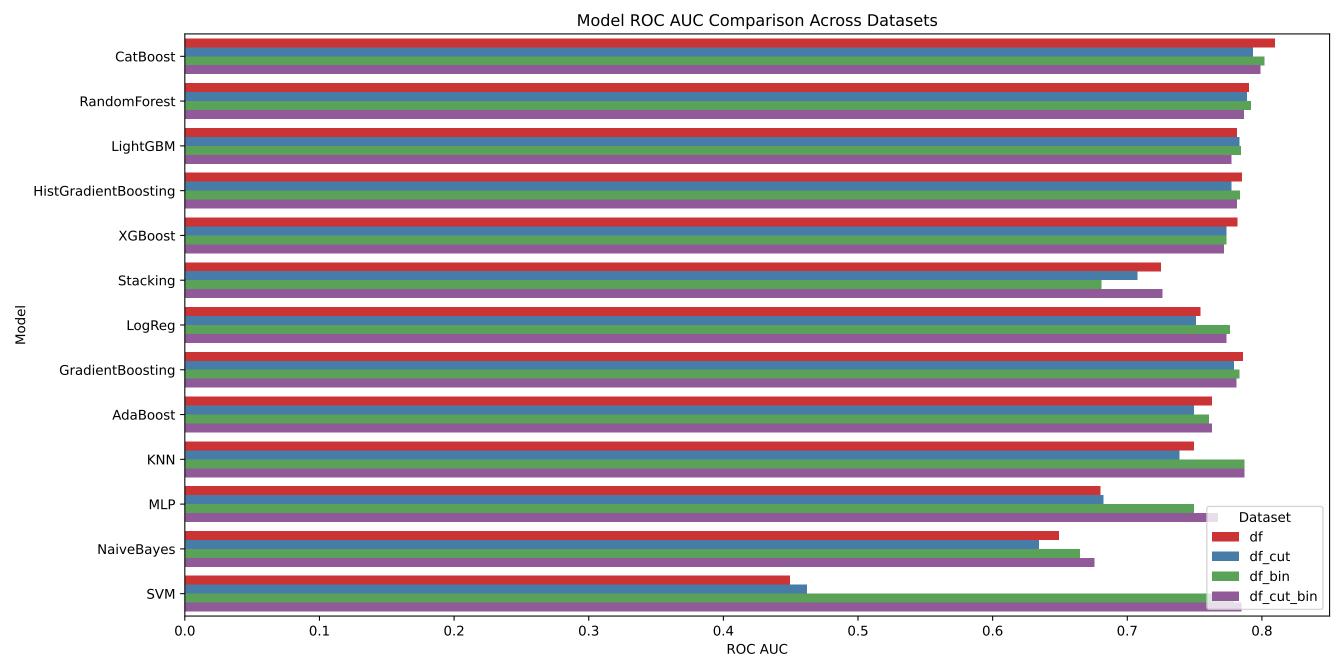
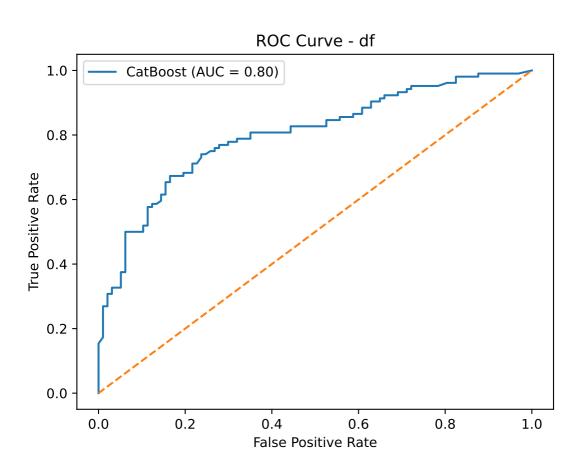
Model F1 Comparison Across Datasets







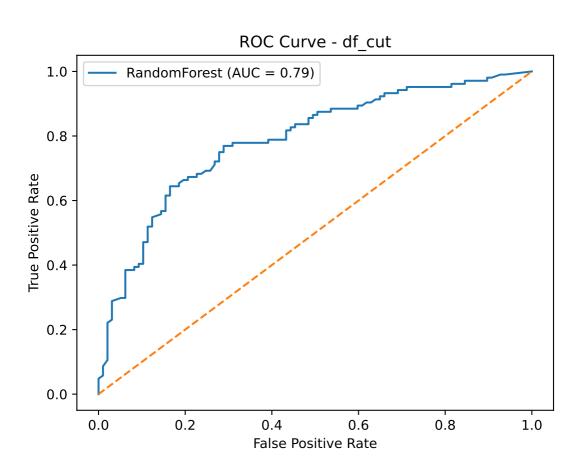
Top 20 Feature Importances - CatBoost (df) BCUT2D_MRLOW -VSA_EState8 -NumHDonors -FpDensityMorgan3 -PEOE_VSA_std -PEOE_VSA9 -SlogP_VSA5 -NHOHCount -MolWt_TPSA_ratio -Feature EState_VSA7 -MaxAbsPartialCharge -PEOE_VSA7 PEOE_VSA_max -NumAromaticCarbocycles -EState_VSA6 -Chi2v -VSA_EState6 -SlogP_VSA6 -FractionCSP3 -NumSaturatedHeterocycles -

3

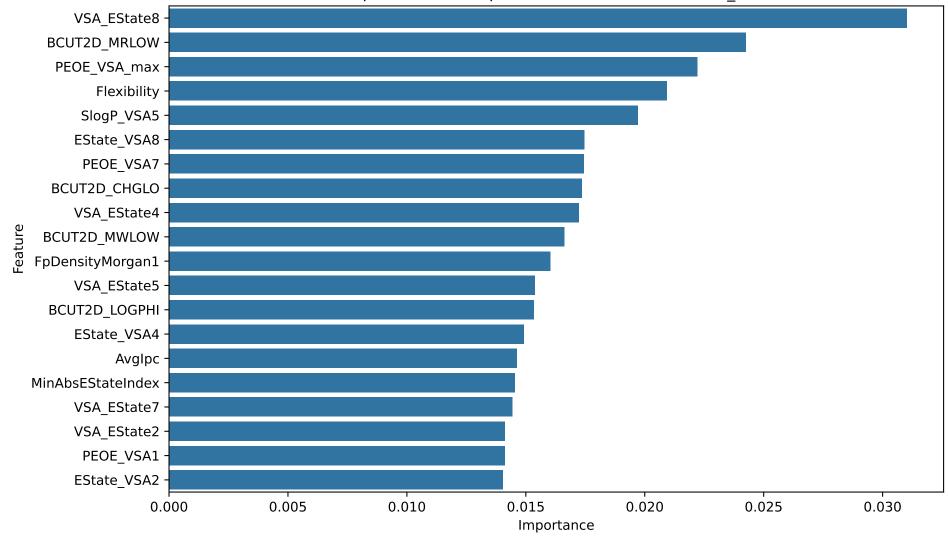
Importance

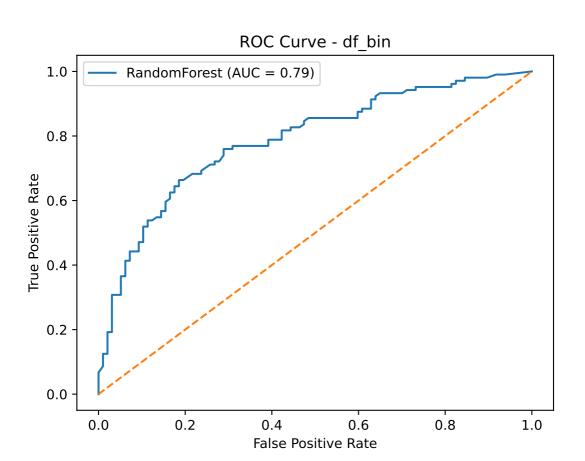
5

6

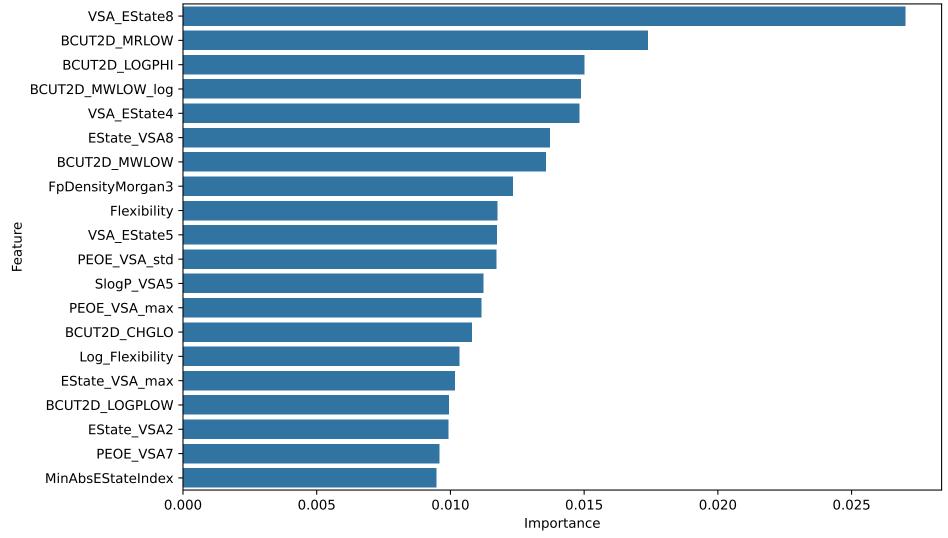


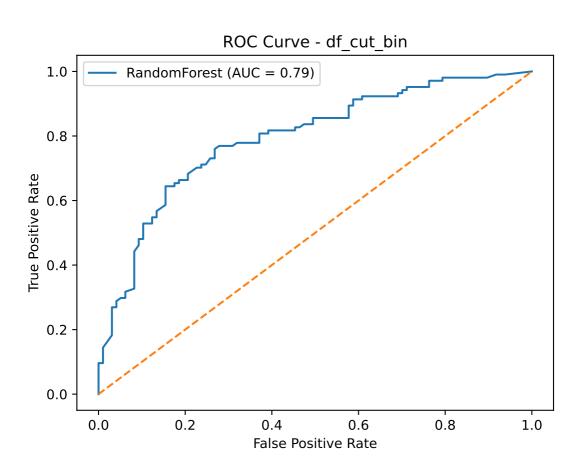
Top 20 Feature Importances - RandomForest (df_cut)





Top 20 Feature Importances - RandomForest (df_bin)





Top 20 Feature Importances - RandomForest (df_cut_bin)

