Results Analysis

The ligand binding predictor was evaluated on ten proteins using a script that compares predicted pocket coordinates against known ligand atom positions. The accuracy metric used was the proportion of predicted pockets that were considered “matched” — those falling within a specific distance of a ligand atom — and an additional metric tracked the average minimum distance from each predicted pocket to the nearest ligand atom.

In all ten test cases, the predictor achieved an accuracy of 0.00, indicating that none of the predicted pockets were classified as matches. For proteins like **1A52**, **1AVW**, **1D86**, and **1HSG**, despite having dozens of known ligand atoms, none of the predicted pockets overlapped spatially with those ligands. The minimum distances between predicted pockets and ligand atoms in these cases ranged from approximately 17 to 35 Å, which is well beyond typical binding site proximity thresholds. For example, 1D86 reported an average minimum distance of 17.3 Å, while 2F5N showed distances as high as 37.7 Å.

In some cases (e.g., 2F5N), the closest predicted pocket was only about 10.7 Å away from the nearest ligand atom, suggesting partial success in localizing binding-relevant regions, albeit not within the matching threshold.

Overall, while the program consistently produced pocket candidates, these did not correlate spatially with true ligand-binding sites, highlighting limitations in the current implementation.

PDB ID	Ligand Atoms Found	Predicted Pockets	Accuracy	Avg. Min Distance (Å)
1A52	46	1	0.00	23.03
1AVW	1	4	0.00	35.44
1D86	31	3	0.00	17.31
1HSG	45	3	0.00	24.02
2F5N	13	5	0.00	26.54

2RH1	213	3	0.00	19.15
3PTB	10	4	0.00	22.31
4Q21	29	1	0.00	27.91

Discussion and Interpretation

The uniformly low performance of the ligand-binding predictor suggests there are significant gaps between the predicted surface pockets and actual ligand positions. Several key factors likely contribute to this discrepancy.

First, the depth-based filtering and surface geometry metrics used to identify candidate pockets may not be sufficient on their own to localize biologically meaningful binding sites. Many true ligand binding sites, while partially buried, are not among the deepest geometric pockets on the protein surface. If a pipeline strongly favors high surface depth as a proxy for binding potential, it may systematically exclude viable but shallower sites.

Second, the clustering approach (DBSCAN) is highly sensitive to parameter selection. Suboptimal values for the ‘eps’ and ‘min_samples’ parameters could result in either over-clustering (splitting true binding sites into multiple incomplete fragments) or under-clustering (merging unrelated surface features into false pockets). Without tuning these parameters based on a training set or representative examples, the algorithm may not generalize well across proteins with varying topologies.

Another key limitation is that the evaluation currently relies on centroid distances between predicted pockets and ligand atoms. This metric is useful but coarse; it fails to capture whether the pocket actually overlaps with the ligand volume. A predicted pocket could be reasonably close (e.g. within 10–15 Å) and still not be considered a match under strict criteria. Including volumetric overlap, surface-contact scoring, or interaction-based features could refine the match definition and yield more informative performance metrics.

In addition, protein flexibility is not accounted for in the current method. Ligands can sometimes induce subtle conformational changes in binding regions. If the pipeline is applied to apo (ligand-free) conformations, pockets may appear structurally different than they are in the ligand-bound state, further complicating detection.

Suggestions for Improvement

- **Parameter Optimization/Machine Learning**: Systematically tune DBSCAN and depth thresholds using cross-validation on a curated training set.

- **Refined Evaluation Metrics**: Consider overlap-based scoring, or count ligand atoms that fall within pocket boundaries rather than just relying on centroid distances.

- **Hybrid Scoring**: Combine geometric features (e.g., surface depth, pocket volume) with biochemical or evolutionary data, such as residue conservation or hydrophobicity.

- **Benchmark Comparison**: Run the same test set through existing tools like Fpocket or CASTp to benchmark your method's recall and precision.

- **Ligand-Aware Filtering**: During development, include known ligand positions in the analysis pipeline to help tune the filtering of spurious predictions.

- **Conformational Considerations**: Where possible, use holo (ligand-bound) structures for training and evaluation, or account for dynamics in pocket detection.