

Final Project

Analysis of Synapse Segregation in Neuronal Branches

Ghaida Zoubi & Donia shebrawi

MSc Statistics - Data Science

University of Haifa

October 18, 2025

Abstract

This study investigates whether synapse placement along neuronal branches occurs randomly or follows structured organizational patterns. Using a large dataset of over 9 million valid branch splits from approximately 130,000 neurons, we combined classical statistical tests, null simulations, and machine learning approaches to assess the degree of synapse segregation. Fisher’s exact test identified significant branch splits, and a binomial test confirmed that the observed proportion of significant splits (11.3%) was substantially higher than the 5% expected by chance. False Discovery Rate (FDR) correction and null simulations verified that this result was not driven by statistical bias. Depth-dependent analyses demonstrated that the proportion of significant splits decreased with increasing dendritic depth, reflecting organized structural patterns along the neuronal tree. Random Forest and Generalized Linear Model (GLM) analyses also showed that both structural and connectivity features—particularly synapse counts, depth, and spatial attributes—were strong predictors of split significance. Together, these results provide converging evidence that synaptic distribution is not purely random but instead follows structured, neuron-specific patterns associated with morphology and functional organization.

1 Introduction

Neurons communicate through synapses, whose distribution across axonal and dendritic branches shapes the computational properties of neural circuits. Understanding whether synapse placement occurs randomly or follows structured patterns is important for revealing principles of neuronal organization and information processing.

Previous studies have suggested that synaptic distribution may not be uniform, raising the question of whether certain branch points preferentially host synapses, potentially reflecting functional specialization. However, characterizing such patterns requires rigorous statistical analysis that accounts for neuron-specific and branch-level variability.

The objective of this project is to investigate synapse segregation at neuronal branch points, assessing if observed distributions deviate from chance. Using Fisher’s exact test, false discovery rate (FDR) correction, machine learning models, and generalized linear modeling (GLM), we aim to assess the degree of randomness and identify structured patterns across neurons and brain areas.

Research question: Are synapses randomly distributed at neuronal branch points (splits), or do they exhibit organized segregation (polarization) between axonal and dendritic branches?

2 Data Description and Pre-processing

2.1 Original Data

The dataset comprises **42,098,838 neuronal splits** from **131,876 unique neurons** and 22 features. Each neuron contributes a variable number of splits, ranging from **4** to **10,439**, with an average of approximately **319 splits per neuron**.

The dataset includes several numeric features such as synapse counts `c1_pre`, `c1_post`, `c2_pre`, and `c2_post`, which exhibit wide variation across branches — in some cases reaching values above 18,000 synapses per split.

The dataset includes several morphological and structural features of neurons. The main variables used in the analysis are the `super_class`, which groups neurons into broad functional categories which was merged from a separate dataset of 133,629 neurons without missing values to assign each neuron to a functional category, and `depth`, representing the number of node transitions from the root within each neuron’s tree structure

And the following is a summary of the variables that we used in our analysis including ML and GLM models:

Table 1: Description of main variables used in the analysis

Variable	Type	Description
<code>super_class</code>	Categorical	Functional category of the neuron
<code>depth</code>	Numeric	Node depth in the neuron tree (jumps from root)
<code>size_nm</code>	Numeric	Size of the neuron
<code>nodes</code>	Numeric	Number of nodes in the neuron structure
<code>cable_length</code>	Numeric	Total length of the neuron
<code>SI</code>	Numeric	Polarization measure (0 = random, 1 = fully polarized)
<code>out_synapses_count</code>	Numeric	Number of presynaptic sites
<code>in_synapses_count</code>	Numeric	Number of postsynaptic sites
<code>c1_pre</code>	Numeric	Pre-synapses in branch 1
<code>c1_post</code>	Numeric	Post-synapses in branch 1
<code>c2_pre</code>	Numeric	Pre-synapses in branch 2
<code>c2_post</code>	Numeric	Post-synapses in branch 2
<code>x</code>	Numeric	X coordinate in space
<code>y</code>	Numeric	Y coordinate in space
<code>z</code>	Numeric	Z coordinate in space
<code>max_dist_from_leaf</code>	Numeric	Max hops from node to leaf
<code>min_dist_from_leaf</code>	Numeric	Min hops from node to leaf
<code>down_stream_splits</code>	Numeric	Number of splits downstream of node

2.2 Data Cleaning and Summary

Initial data cleaning and merging was performed by removing rows with missing values, which ensures that all analyzed splits had complete information for synapse counts and geometric properties. After this filtering step, the dataset contained **9,341,305 valid splits** from **130,229 unique neurons**, with the number of splits per neuron ranging from **1** to **3,101** (mean ≈ 71.7 splits per neuron).

Descriptive statistics of our main variables that we used showed large variability in synapse counts (`c1_pre`, `c1_post`, `c2_pre`, `c2_post`) and node depth (`depth`). This highlights the need to account for differences between neurons and branches in the subsequent analyses.

To ensure the reliability of the precomputed p-values, Fisher’s exact test was recalculated for each split using the original synaptic counts (`c1_pre`, `c1_post`, `c2_pre`, `c2_post`). The recalculated p-values were then compared against those stored in the dataset to confirm computational consistency, no substantial discrepancies were observed, confirming that the original p-values were accurate and suitable for subsequent analyses.

A new binary column `significant` was added to indicate whether the Fisher exact test for each split was significant at $\alpha = 0.05$.

In the following analysis we also computed the proportion of significant splits per branch length by grouping the data on `depth` after adding a relevant variable which will be explained in the following sections, we then calculated the proportion of significant splits at each `depth` level.

Based on biological theory, which assumes that the first split perfectly distributes axons and dendrites, we removed the first split of each neuron.

Most of the subsequent analyses were conducted on the dataset excluding these first splits; The data with no first splits includes 9211076 splits and 24 variables (with the new variables that we created), in other words 127898 neurons, in this data we have less neurons because of removing the neurons that includes only one split.

3 Methods

3.1 Significance Testing of Synapse Segregation

To determine if the overall proportion of significant branch splits exceeded what would be expected by random chance, we applied a binomial test. Each split was labeled as significant when its Fisher’s exact test p -value was lower than $\alpha = 0.05$. The observed number of significant splits (k) was compared to the total number of valid splits (n) under the null hypothesis that $p \leq 0.05$, corresponding to a random distribution of synapses. The binomial test was conducted using a one-sided alternative hypothesis ($H_1 : p > 0.05$), evaluating if the empirical rate of significant splits was greater than expected by chance. Additionally, a sensitivity analysis excluding initial splits was used in our Binomial analysis in order to ensure our results were not dominated by the first split of each neuron (which defined by smallest depth), we repeated our analysis after their removal.

3.2 Null simulation of Fisher’s test

To verify that the observed rate of significant splits was not inflated by statistical bias, we simulated data under the null hypothesis of random synapse distribution. For each of 10 iterations, 5000 branch splits were sampled and their synaptic counts were randomly reassigned between branches assuming a 50:50 probability using `np.random.binomial`. Fisher’s exact test was applied to each randomized table, and the proportion of $p < 0.05$ results was recorded.

That simulation checks that the proportion of significant splits when taking random simulated data is less than expected by chance.

3.3 Additional Robustness Checks - FDR

To another validate the robustness of the findings we have got, we decided to conduct complementary analyses on the dataset.

First of all, we applied a False Discovery Rate (FDR) correction by using the Benjamini–Hochberg procedure to account for multiple comparisons.

Second, We measured how much the observed proportion of significant splits differed from the random expectation using Cohen’s h , which quantifies the size of this difference. Cohen’s h was computed using the following formula:

$$h = 2 \arcsin(\sqrt{p_1}) - 2 \arcsin(\sqrt{p_2})$$

where p_1 and p_2 represent the observed and expected proportions of significant splits, respectively.

Collectively, these two validation steps are crucial and They ensure our conclusions reflect a true biological signal, not just random noise or statistical artifacts from multiple testing.

3.4 Dependence of Significant Splits Within Neurons

We used a chi-squared test of independence To evaluate whether significant synaptic splits occur independently or tend to cluster within specific neurons, Each split was labeled as significant if its Fisher’s exact test p -value was below $\alpha = 0.05$. For each neuron, we counted the number of significant and non-significant splits and applied a chi-squared test to determine if the distribution of significant splits was independent across neurons.

The hypotheses for the chi-squared test were formulated as follows:

- Null hypothesis (H_0): Split significance is independent of the neuron.
- Alternative hypothesis (H_1): Split significance depends on the neuron.

3.5 Depth Group Analysis and Significance Testing

To evaluate whether the significance of synaptic splits depends on dendritic depth, all branch splits were grouped into five depth intervals (1–5, 6–10, 11–15, 16–20, and 21+ units, with the final bin including all larger values).

Depth values were binned using standard Python tools (`pd.cut`) to create categorical depth groups for subsequent analyses..

For each depth group, the number and proportion of significant splits were calculated using Fisher’s exact test ($\alpha = 0.05$). Building on the global binomial test described above which assess whether the overall proportion of significant splits exceeded the expected 5% under the null hypothesis ($H_0 = 0.05$). We performed, binomial tests were conducted separately within each depth group to examine whether significance varied across dendritic depth.

To account for neuron-level aggregation, pre- and post-synaptic counts (`c1_pre`, `c1_post`, `c2_pre`, `c2_post`) were summed for each neuron within each depth group. Fisher’s exact test was then recalculated per neuron per depth group to determine split-level significance while controlling for intra-neuron dependence.

To correct for multiple comparisons, p-values obtained from per-neuron Fisher tests were adjusted using the Benjamini–Hochberg False Discovery Rate (FDR) procedure. For each depth group, we reported both the raw and FDR-corrected number and proportion of significant splits, along with corresponding binomial p-values.

This workflow allowed us to assess depth-dependent patterns of synapse segregation, account for multiple comparisons, and control for potential neuron-level dependence in the dataset.

3.6 Superclass Analysis

We analyzed neurons organized by their `super_class` which contains nine categories: *optic*, *central*, *sensory*, *visual_projection*, *visual_centrifugal*, *ascending*, *descending*, *motor*, and *endocrine*, to explore the distribution of significant splits. The purpose of this analysis was to assess whether synapse segregation occurred systematically across distinct neuron categories. For each superclass, we calculated the proportion of significant splits among all valid branch splits and compared these proportions across classes.

Extreme neurons with 0% or 100% significant splits were identified to characterize neurons with fully random or fully segregated splits, Before doing all the analysis we started with showing that each neuron was assigned to a single superclass. For each neuron, the proportion of significant splits is calculated as:

$$\text{Proportion of significant splits for neuron } i = \frac{\text{Number of significant splits in neuron } i}{\text{Total number of splits in neuron } i}$$

3.7 Machine Learning Analysis

To predict whether a neuronal branch split exhibits statistically significant synapse segregation, we implement a supervised machine learning by using a Random Forest classifier. The model was built with 200 decision trees (`n_estimators = 200`) and a maximum depth of 15 (`max_depth = 15`) and class weights set to `balanced` to deal with put imbalanced data, using the `scikit-learn` library in Python. The input features included

quantitative features (e.g., `c1_pre`, `c1_post`, `c2_pre`, `c2_post`, `depth`, and structural metrics) and categorical variables (`super_class`, `depth_group`, `area_nm`) were used as input, with categorical/morphological features one-hot encoded. The target variable indicated whether each branch split exhibited statistically significant synapse segregation, as determined by Fisher’s exact test.

Our dataset was randomly divided into training (80%) and testing (20%) subsets. Model performance was evaluated using accuracy, precision, recall, and F1-score, confusion matrix, and ROC-AUC. Furthermore, feature importance values were extracted to identify the most influential predictors contributing to model performance.

The measure that will be reported in the results to show the most influential predictors for significant synapse segregation is **Gini impurity**:

$$G = 1 - \sum_{i=1}^C p_i^2$$

where C is the number of classes and p_i is the proportion of samples of class i in the node. Lower Gini values correspond to purer nodes.

Lower Gini values means purer nodes, meaning that minimizing the Gini impurity leads to better splits and thus better predictive power for the model.

3.8 Generalized Linear Model (GLM)

To complement the Random Forest analysis and provide inferential evidence, we implemented a Generalized Linear Model (GLM) with a binomial family to model the probability that a neuronal split is statistically significant. To reduce computational load and address class imbalance, we randomly sampled an equal number of neurons with significant and non-significant splits (up to 1,000 neurons per group), retaining all splits from the selected neurons.

The model contains a combination of structural, connectivity, and spatial features, such as neuron size, number of nodes, cable length, depth, maximum and minimum distances from the leaf, downstream splits, signal integration (SI), pre- and post-synaptic counts on both branches, as well as the total number of pre- and post-synapses and 3D coordinates (x, y, z). Categorical variables, including super-class and depth group, were also incorporated. The outcome variable was binary, indicating whether a split was significant.

To account for correlation among splits within the same neuron, cluster-robust standard errors were applied with neurons as the clustering unit. The GLM was fitted using the `statsmodels` library in Python with a binomial family and cluster-robust covariance estimation. This modeling approach allowed us to quantify the association between structural and connectivity features and split significance while controlling for intra-neuron dependence, providing rigorous statistical evidence to test our hypothesis that synapse placement along neuronal branches follows structured, neuron-specific patterns rather than occurring randomly.

4 Results

4.1 Significance Testing of Synapse Segregation

Analysis of all valid branch splits showed that approximately 11.32% were statistically significant according to Fisher’s exact test ($p < 0.05$).

A sensitivity analysis excluding the first split of each neuron showing 11.22% of the splits were significant.

4.2 Null simulation of Fisher’s test

Table 2: Proportion of significant splits under the null hypothesis (simulated data).

Iteration	Proportion significant (%)
1	1.64
2	2.04
3	2.02
4	1.58
5	1.84
6	1.84
7	2.08
8	1.94
9	2.24
10	1.94
Average	1.92

As we can see the simulation shows that when synaptic counts are randomly reassigned under the null hypothesis, the proportion of significant splits is consistently below the expected 5%.

4.3 Additional Robustness Checks - FDR

Table 3: FDR-Corrected Significant Splits and Effect Size (Cohen’s h)

Metric	Value
Number of significant splits (FDR)	242,802
Proportion of significant splits (FDR)	0.0264
Cohen’s h (observed vs expected)	0.2322

After applying the FDR correction, approximately 2.64% of all branch splits (~242800 in total) remained significant, and the Cohen’s effect value is 0.2322.

4.4 Chi-Squared Test of Split Dependence Across Neurons

Table 4: Chi-squared Test for Independence of Split Significance Across Neurons

Test	Chi ² Statistic	df	p-value
Chi-squared test	479925.57	127897	< 0.001

In the chi-squared test the value of the statistic is 479925.57 with 127897 degrees of freedom and p-value < 0.001 which indicated a significant association between neurons and split significance.

4.5 Depth Group Analysis and Significance Testing

Table 5: Split-level significance per depth group (Fisher + Binomial + FDR)

Depth Group	Total Splits	Significant	Prop. Sig.	Binom p	FDR Sig.	Prop. FDR	Binom FDR p
1-5	1,757,473	284,767	16.20%	0	75,307	4.28%	1
6-10	3,075,796	362,110	11.77%	0	78,298	2.55%	1
11-15	1,994,179	202,280	10.14%	0	44,195	2.22%	1
16-20	1,063,370	93,200	8.76%	0	21,620	2.03%	1
21+	1,320,258	91,342	6.92%	0	23,382	1.77%	1

After grouping the data by depth group we can see that the proportion of significant splits in the binomial test ranged from 6.92% to 16.2%, all with p-values < 0.01. However after correcting for multiple comparisons using FDR, the range of the proportions decreased to 1.77% - 4.28% all remaining under 5%.

4.6 Superclass Analysis

We analyzed neurons classified by their *super_class* in order to examine the distribution of significant dendritic/axonal splits. Each neuron was assigned to exactly one *super_class*, as we can see in the following table:

Table 6: Neuron assignment to *super_class*

Total neurons	127,898
Neurons with more than one <i>super_class</i>	0
Rows with <i>super_class</i>	9,211,076
Rows with missing <i>super_class</i>	0

Table 7: Descriptive statistics per **super_class** (split-level)

Super Class	Total Splits	Num Neurons	Prop. Sig.	Prop. Total Splits
central	4,125,168	32,262	11.35%	44.78%
optic	3,504,543	76,290	12.72%	38.05%
visual_projection	971,926	8,030	8.14%	10.55%
descending	372,571	1,236	5.08%	4.04%
visual_centrifugal	94,765	486	9.43%	1.03%
ascending	55,000	1,809	14.91%	0.60%
motor	42,568	104	2.21%	0.46%
sensory	40,537	7,625	8.59%	0.44%
endocrine	3,998	56	1.28%	0.04%

In table 7 we can see the 9 classes of **super_class** variable with the proportion of significant splits for each level, the total number of splits, the number of neurons per class, and the proportion of total splits contributed by each class.

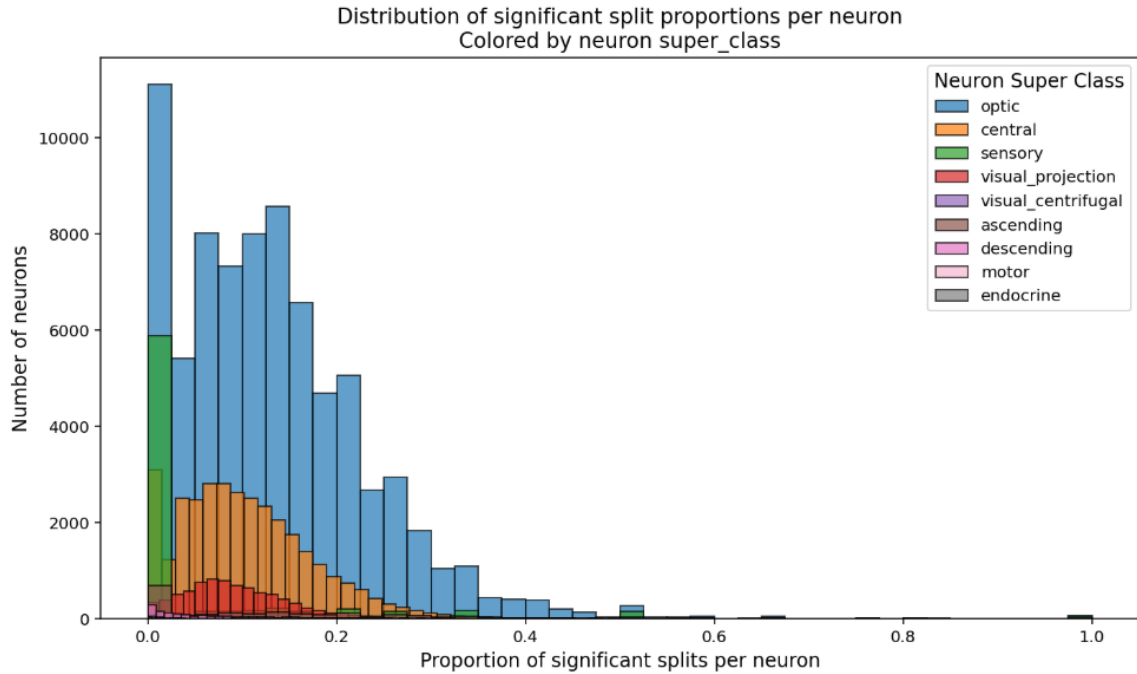


Figure 1: Distribution of significant dendritic/axonal splits across neurons grouped by *super_class*.

Figure 1 shows the distribution of significant split proportions per neuron, with colors indicating the **super_class** levels.

Table 8: Extreme neurons based on proportion of significant splits per **super_class**

Extreme	Total Neurons	Top Super Classes	Range of Splits per Neuron	Min Splits	Max Splits
0% significant splits	20,394	optic, sensory	1–377	1	377
100% significant splits	109	optic, sensory	1–4	1	4

Table 8 shows descriptive statistics for the extreme neurons based on proportion of significant splits per **super_class**, there are 20,394 neurons with no significant splits,

with the number of splits ranging from 1 to 377. In contrast, 109 neurons have all splits significant, with the number of splits ranging from 1 to 4.

4.7 Random Forest Classification -Machine Learning Analysis

4.7.1 Confusion Matrix

Table 9: Confusion Matrix for ML Model Predicting Split Significance

	Predicted 0	Predicted 1
Actual 0	1,095,461	540,015
Actual 1	70,397	136,343

Table 9 shows the number of true positives, true negatives, false positives, and false negatives for the ML model

4.7.2 Classification Metrics

Table 10: Classification Report for ML Model

Class	Precision	Recall	F1-score	Support
0	0.9396	0.6698	0.7821	1,635,476
1	0.2016	0.6595	0.3088	206,740
Accuracy	0.6687			
Macro Avg	0.5706	0.6647	0.5454	1,842,216
Weighted Avg	0.8568	0.6687	0.7290	1,842,216

In Table 10 we can see the strength of the Random forest model, the accuracy of the model is 0.6687, the **macro average** reports the mean precision, recall, and F1-score across both classes (Class 0: non-significant splits, Class 1: significant splits) treating them equally, while the **weighted average** accounts for the number of samples in each class, giving more weight to the majority class (Class 0) in the metrics.

4.7.3 ROC-AUC

Table 11: ROC-AUC Score for ML Model

Metric	Value
ROC-AUC	0.7328

In Table 11, the ROC-AUC score of $0.7328 > 0.5$ shows better ability of the model to predict the significance of the split than random guessing.

4.7.4 Top 10 Feature Importance

Table 12: Top 10 Features by Importance in ML Model

Feature	Importance
c2_pre	0.0737
c1_pre	0.0615
c1_post	0.0596
min_dist_from_leaf	0.0560
SI	0.0474
depth	0.0430
max_dist_from_leaf	0.0395
depth_group_1-5	0.0363
c2_post	0.0347
down_stream_splits	0.0342

The Table above shows the top 10 most important features identified by the Random Forest model in predicting whether a neuronal split is significant. The importance values correspond to **Gini importance**, as described previously, which measures how much each feature contributes to reducing impurity in the decision trees.

4.8 Generalized Linear Model (GLM) Results

4.8.1 GLM Table for Split Significance

The table below shows only statistically significant predictors ($p < 0.05$) from the balanced sample.

Positive coefficients mean that as the value of the feature increases, the probability that a split is significant also increases (higher likelihood), while negative coefficients mean that higher values of that feature are associated with a lower probability of a split being significant.

Table 13: GLM results for predicting split significance (balanced sample). Only significant variables are shown.

Variable	Coefficient	Std. Error	z-value	p-value
Intercept	-1.981	0.212	-9.353	0.000
C(super_class)[T.endocrine]	-2.534	0.232	-10.922	0.000
C(super_class)[T.motor]	-2.649	0.218	-12.146	0.000
C(depth_group)[T.6-10]	-0.226	0.030	-7.535	0.000
C(depth_group)[T.11-15]	-0.344	0.038	-8.974	0.000
C(depth_group)[T.16-20]	-0.432	0.054	-7.953	0.000
C(depth_group)[T.21+]	-0.525	0.067	-7.856	0.000
nodes	0.0001	3.71e-05	3.058	0.002
SI	-1.597	0.123	-12.987	0.000
cable_length	-5.85e-07	2.00e-07	-2.924	0.003
c1_pre	0.0020	0.000	7.469	0.000
c1_post	0.0014	0.001	2.413	0.016
c2_pre	0.0021	0.000	5.572	0.000
c2_post	0.0011	0.001	2.022	0.043
out_synapses_count	0.0002	3.56e-05	5.097	0.000
max_dist_from_leaf	0.0231	0.005	4.268	0.000
min_dist_from_leaf	0.5307	0.034	15.418	0.000
down_stream_splits	-0.0093	0.003	-2.704	0.007

5 Discussion

As shown in the results, before grouping the depth variable, 11.32% of splits were significant, which is above the 5% expected by random chance. This was confirmed by a Fisher test on simulated data, which consistently showed proportions below 5%. Following biological theory, which suggests that the first split separates axons and dendrites almost perfectly, we continued our analysis after removing the first split from each neuron. After this adjustment, the binomial test still identified 11.22% significant splits, slightly lower than before but still above 5%. However, after FDR correction, the proportion dropped to 2.64%, below the threshold expected under the null hypothesis, indicating that, when accounting for multiple testing, split significance appears largely random. Cohen’s effect size of 0.2322 further supports this, showing only a small effect.

The occurrence of significant splits also depends on the neuron itself. The chi-square test showed a very high statistic, leading to rejection of the null hypothesis that split significance is independent of the neuron, indicating a clear neuron-specific effect. This dependence is also evident when neurons are grouped by super-class. Descriptive statistics reveal that most classes have proportions of significant splits above 5%, except for the motor (2.21%) and endocrine (1.28%) classes, which have lower proportions. This difference is not simply due to the total number of splits, as motor and sensory classes have similar numbers of splits (42568, 40537), yet the sensory class shows a higher proportion of significant splits (8.59%). Similarly, descending class has a higher number of splits (372571) than sensory, yet a lower proportion of significant splits (5.08%). These patterns, illustrated in Figure 1, show an asymmetric distribution with a rightward tail across the classes.

We then examined extreme values in the distributions: neurons with 0% and 100% significant splits. Neurons with no significant splits numbered 20,394, with splits ranging from 1 to 377, while neurons with all splits significant were only 109, with 1–4 splits each. Interestingly, optic and sensory classes are overrepresented at both extremes, suggesting unique and shared properties in these neuron types.

To address depth-related dependencies, we created a depth group variable, categorizing all branch splits into five intervals (1–5, 6–10, 11–15, 16–20, and 21+ units). Applying the binomial test and FDR correction within these groups, we found that all depth groups had proportions above 5% in the binomial test, while FDR-corrected proportions were below 5%, except for the first group (depth 1–5), which had the highest corrected proportion at 4.28%. This supports the idea that shallow splits, closer to the root, are more likely to be significant, consistent with theoretical expectations about the first split.

Next, we applied a Random Forest model, which performed well with an accuracy of 0.6687 and a ROC-AUC score of 0.7328, indicating reasonable predictive power. Feature importance analysis showed that the most influential variables for predicting split significance included `c2_pre`, `c1_pre`, `c1_post`, minimum distance from leaf, SI, depth, maximum distance from leaf, depth group 1–5, `c2_post`, and downstream splits. Notably, neurons in the depth 1–5 group played a major role in determining split significance.

Finally, we fitted a GLM with a binomial family on a balanced sample. Positive coefficients indicate that higher values of a feature increase the probability of a split being significant, while negative coefficients indicate the opposite. Strong negative associations were observed for SI, motor and endocrine super-classes, deeper depth groups, downstream splits, and cable length, suggesting that higher values reduce the likelihood of a split being significant. Positive associations included minimum distance from leaf, `c2_pre`,

c1_pre, maximum distance from leaf, nodes, c1_post, c2_post, and out_synapses_count, highlighting that branch position and pre-synaptic counts are key contributors to split significance.

We suggest Future work which could incorporate additional structural or functional variables to further investigate synapse organization.

6 Conclusion

This project examined whether synapse placement along neuronal branches occurs randomly or follows structured, neuron-specific patterns that reflect functional organization. The analyses revealed non-random, depth-dependent segregation of synapses, suggesting that certain neuronal regions exhibit structured synaptic organization. The Random Forest model effectively distinguished between random and structured distributions, highlighting the predictive importance of structural and connectivity features such as synapse counts and spatial attributes. Complementary results from the Generalized Linear Model (GLM) reinforced these findings, demonstrating consistent relationships between synaptic patterns and neuronal properties. Overall, the findings indicate that synaptic distribution is not purely stochastic but exhibits structured tendencies linked to neuron morphology and depth, offering insights into the organizational principles of neural connectivity.

References

- [1] Hawkins, J., & Ahmad, S. (2016). Why neurons have thousands of synapses: A theory of sequence memory in neocortex. *Frontiers in Neural Circuits*, 10, 23. <https://doi.org/10.3389/fncir.2016.00023>
- [2] Goaillard, J.-M., et al. (2020). Diversity of axonal and dendritic contributions to neuronal output. *Nature Reviews Neuroscience*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6987044/>
- [3] Ujfalussy, B. B., & Makara, J. K. (2020). Impact of functional synapse clusters on neuronal response selectivity. *Nature Communications*, 11(1), 1413. <https://doi.org/10.1038/s41467-020-15147-6>