QUANTUM WOLF

DATA INTELLIGENCE AND RESEARCH LAB

DATE:10-03-2025

NAME:ABDUL ZUHAIL M

Smart Dust Networks: Packet Loss Minimization in Submillimeter Sensor Communication

1. Problem Statement

Smart Dust Networks utilize submillimeter-scale wireless sensors for environmental data collection. However, dense deployments result in high signal interference and packet loss of up to 40%, limiting data reliability. This problem significantly affects applications such as precision agriculture and environmental monitoring, where real-time data accuracy is crucial.

2. Introduction

The primary challenge in Smart Dust Networks is managing efficient communication between motes while mitigating signal loss. Conventional routing methods fail to adapt dynamically to interference variations, leading to frequent data retransmissions and inefficient power consumption. This project focuses on minimizing packet loss and improving routing efficiency using AI-driven techniques within Orange, a data visualization and analysis tool. The dataset used in this project is generated by ChatGPT to simulate real-world sensor network conditions.

3. Solution Approach

This project implements a data-driven workflow in Orange to optimize signal routing. The methodology involves:

- **Data Preprocessing & Feature Engineering** – Ensuring high-quality input for modeling.
- **Feature Selection & Dimensionality Reduction** – Identifying key attributes affecting routing.
- **Clustering for Network Analysis** – Understanding interference zones.
- **Machine Learning-Based Routing Decisions** – Implementing a decision tree model for optimized routing.

- **Python-Based Adaptive Routing** – Fine-tuning routing paths dynamically using interference predictions.

4. Smart Dust Routing Optimization - Orange Workflow Implementation

Step 1: Data Import

Objective: Load the dataset containing sensor attributes (e.g., signal strength, packet loss, humidity, temperature, and interference levels). The dataset used in this project is generated by ChatGPT.

Widgets Used:

- File (Loads the dataset into Orange.)
- Data Table (Displays raw data for verification.)

Output:

	Mote_ID	X_Coordinate	Y_Coordinate	Signal_Strength (dBm)	Humidity (%)	Temperature (°C)	Packet_Loss (%)	Interference_Level
1	Mote_1	3.7454	6.42032	-66.9063	48.4468	29.1448	34.945	Medium
2	Mote_2	9.50714	0.8414	-42.9234	53.9295	18.0508	55.419	High
3	Mote_3	7.31994	1.61629	-54.8424	48.8505	26.5258	53.887	High
4	Mote_4	5.98658	8.98554	-45.2063	44.4351	27.1343	57.3222	High
5	Mote_5	1.56019	6.06429	-60.3985	46.0318	23.4826	47.6671	High
6	Mote_6	1.55995	0.0919705	-43.1343	63.0389	29.7289	56.5878	High
7	Mote_7	0.580836	1.01472	-58.3239	50.3167	33.6873	48.5988	High
8	Mote_8	8.66176	6.63502	-69.6749	58.2135	33.5114	38.8171	High
9	Mote_9	6.01115	0.0506158	-42.8385	65.1709	24.0168	52.4967	High
10	Mote_10	7.08073	1.60808	-67.2614	74.5197	17.2648	37.6328	High
11	Mote_11	0.205845	5.48734	-60.4206	41.9656	34.6968	41.7915	High
12	Mote_12	9.6991	6.91895	-41.4981	79.9705	31.778	59.8767	High
13	Mote_13	8.32443	6.51961	-41.4818	71.395	17.4933	59.4406	High
14	Mote_14	2.12339	2.24269	-52.7969	44.088	33.4168	47.5974	High
15	Mote_15	1.81825	7.12179	-51.0449	83.6789	32.3979	56.0109	High
16	Mote_16	1.83405	2.37249	-56.5466	86.0436	25.3768	52.7058	High
17	Mote_17	3.04242	3.254	-61.2037	43.0539	26.8255	40.6021	High
18	Mote_18	5.24756	7.46491	-60.1401	53.8439	22.9801	45.5394	High
19	Mote_19	4.31945	6.49633	-49.8244	80.3101	16.0952	59.1549	High
20	Mote_20	2.91229	8.49223	-47.4288	77.413	21.7039	50.3395	High
21	Mote_21	6.11853	6.57613	-46.2526	49.2261	31.0571	56.9742	High
22	Mote_22	1.39494	5.68309	-46.3115	50.4675	15.0926	52.9735	High
23	Mote_23	2.92145	0.936748	-67.2638	58.5236	21.67	41.9801	High
24	Mote_24	3.66362	3.67716	-55.1674	64.2261	22.9634	54.5432	High
25	Mote_25	4.5607	2.65202	-68.2732	70.9127	25.7479	41.1694	High
26	Mote_26	7.85176	2.4399	-53.5141	58.4457	33.3971	51.228	High
27	Mote_27	1.99674	9.73011	-56.7541	63.1267	21.9269	51.8663	High
28	Mote_28	5.14234	3.93098	-43.3689	77.3735	21.9391	58.4455	High
29	Mote_29	5.92415	8.92047	-59.4725	41.8342	29.75	43.7185	High
30	Mote_30	0.464504	6.31139	-66.488	52.6218	24.0444	41.8012	High
31	Mote_31	6.07545	7.94811	-65.7102	75.6675	19.4921	34.6598	Medium
32	Mote_32	1.70524	5.02637	-47.1547	84.7603	24.0488	55.9627	High
33	Mote_33	0.650516	5.76904	-51.4535	65.5839	17.8171	50.8466	High
34	Mote_34	9.48886	4.92518	-66.9663	66.6057	18.5277	34.2393	Medium

Step 2: Data Preprocessing

2.1. Select Columns

- **Objective:** Retain only relevant features required for routing decisions.
- **Widgets Used:** Select Columns (Filters necessary features.)

2.2. Handle Missing Values

- **Objective:** Ensure no gaps exist in the dataset.

- **Widgets Used:** Impute (Fills missing numerical values with mean, categorical values with mode.)

2.3. Modify Domain

- **Objective:** Ensure proper data types and conversions where necessary.
- **Widgets Used:** Edit Domain (Modifies feature types.)

Step 3: Feature Ranking & Reduction

3.1. Rank Features

- **Objective:** Identify the most significant features influencing routing efficiency.
- **Widgets Used:** Rank (Ranks features based on predictive power.)

Feature Importance Ranking:

1. Signal Strength (dBm)
2. Packet Loss (%)
3. Y Coordinate
4. Humidity (%)
5. Temperature (°C)
6. X Coordinate

3.2. Apply PCA for Dimensionality Reduction

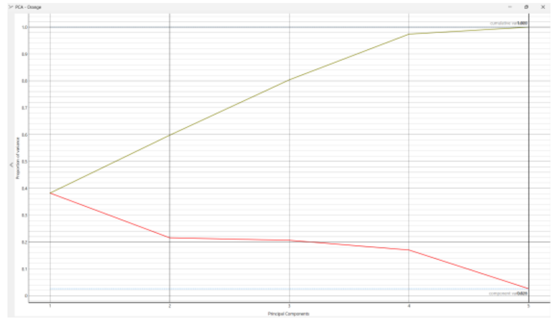
- **Objective:** Reduce redundant data while preserving critical variance.
- **Widgets Used:** PCA (Performs dimensionality reduction.)

PCA Analysis:

- The **variance explained** by the principal components is visualized.
- **Cumulative variance** shows how much of the data variance is retained as more principal components are included.

Output:

		#	Gain ratio	Gini
1	N Signal_Strength (dBm)		0.059	0.018
2	N Packet_Loss (%)		0.059	0.018
3	N Y_Coordinate		0.011	0.003
4	N Humidity (%)		0.009	0.003
5	N Temperature (°C)		0.004	0.001
6	N X_Coordinate		0.002	0.001



Step 4: Clustering & Routing Analysis

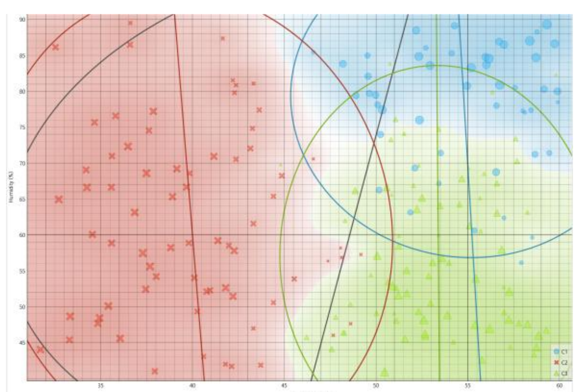
4.1. Apply k-Means Clustering

- **Objective:** Categorize nodes based on network behavior and interference levels.
- **Widgets Used:** k-Means (Performs clustering based on similarity.)

4.2. Visualize Clusters

- **Objective:** Gain insights into cluster distribution.
- **Widgets Used:** Scatter Plot (Plots clustering results.)
- **Box Plot (Analyzes variability in signal parameters.)**

Output:



Step 5: Machine Learning Model for Routing Decisions

5.1. Train a Decision Tree Model

- **Objective:** Develop a predictive model to classify optimal routing decisions.
- **Widgets Used:** Tree (Trains a decision tree classifier.)

5.2. Evaluate Model Performance

- **Objective:** Validate the model’s accuracy and reliability.
- **Widgets Used:** Test & Score (Measures model performance.)
- **Confusion Matrix (Visualizes misclassifications.)**

Confusion Matrix Analysis: The confusion matrix illustrates the model's classification performance. In this case:

- **189 high-signal packets** were correctly classified as **high**.
- **10 medium-signal packets** were correctly classified as **medium**.
- **1 medium-signal packet** was misclassified as **high**.
- **No high-signal packets** were misclassified as medium.

This indicates a highly accurate model, with minimal misclassification errors.

Output:

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.955	0.995	0.995	0.995	0.995	0.951

		Predicted		
		High	Medium	Σ
Actual	High	189	0	189
	Medium	1	10	11
	Σ	190	10	200

Step 6: Python Script for Custom Routing Optimization

6.1. Append Predictions to the Dataset

- **Objective:** Integrate optimized routing decisions.
- **Widgets Used:** Predictions (Adds routing decision predictions.)

6.2. Implement Python-Based Adaptive Routing

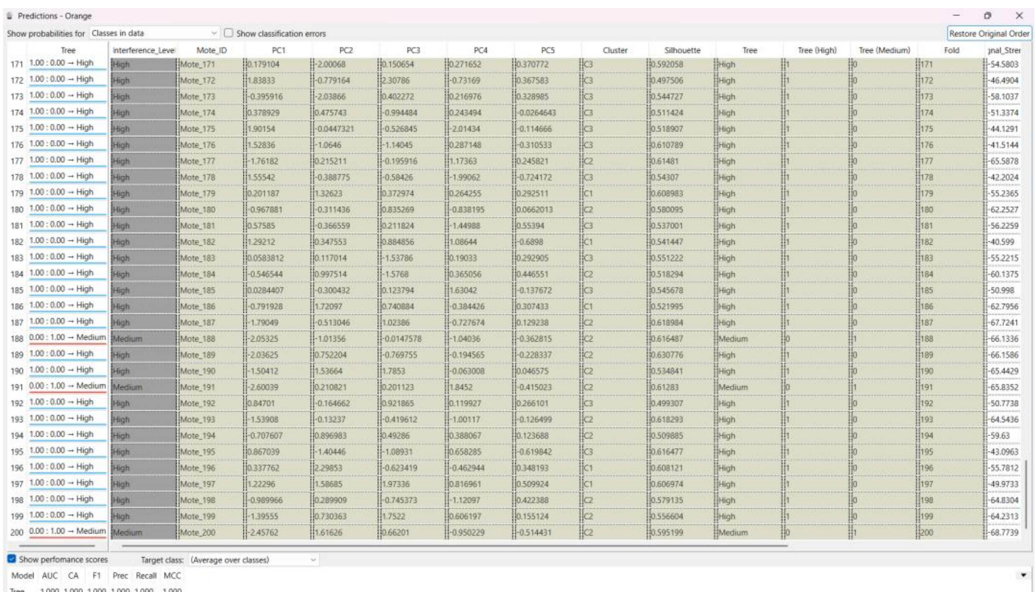
- **Objective:** Fine-tune routing logic based on real-time interference variations
- **Widgets Used:** Python Script (Executes custom routing algorithms.)

6.3. Save Optimized Routing Table

- **Objective:** Store final routing decisions for deployment.
- **Widgets Used:** Data Table (Displays optimized routing data.)
- **Save Data (Stores processed data.)**

Output:

Precision:


Predictions - Orange

Show probabilities for Classes in data | Show classification errors | Restore Original Order

	Tree	Interference_Leve	Mote_ID	PC1	PC2	PC3	PC4	PC5	Cluster	Silhouette	Tree	Tree (High)	Tree (Medium)	Fold	gnal_Strei
171	1.00 : 0.00 → High	High	Mote_171	0.179104	-2.00068	0.150654	0.271652	0.370772	C3	0.592058	High	1	0	171	-54.5803
172	1.00 : 0.00 → High	High	Mote_172	1.83833	-0.779164	2.30786	-0.73169	0.367583	C3	0.497506	High	1	0	172	-46.4904
173	1.00 : 0.00 → High	High	Mote_173	-0.395916	-2.03866	0.402272	0.216976	0.328985	C3	0.544727	High	1	0	173	-58.1037
174	1.00 : 0.00 → High	High	Mote_174	0.378929	0.475743	-0.994484	0.243494	-0.0264643	C3	0.511424	High	1	0	174	-51.3374
175	1.00 : 0.00 → High	High	Mote_175	1.90154	-0.0447321	-0.526845	-2.01434	-0.114666	C3	0.518907	High	1	0	175	-44.1291
176	1.00 : 0.00 → High	High	Mote_176	1.52836	-1.0646	-1.14045	0.287148	-0.310533	C3	0.610789	High	1	0	176	-41.5144
177	1.00 : 0.00 → High	High	Mote_177	-1.76182	0.215211	-0.195916	1.17363	0.245821	C2	0.61481	High	1	0	177	-65.5878
178	1.00 : 0.00 → High	High	Mote_178	1.55542	-0.388775	-0.58426	-1.99062	-0.724172	C3	0.54307	High	1	0	178	-42.2024
179	1.00 : 0.00 → High	High	Mote_179	0.201187	1.32623	0.372974	0.264255	0.292511	C1	0.608983	High	1	0	179	-55.2365
180	1.00 : 0.00 → High	High	Mote_180	-0.967881	-0.311436	0.835269	-0.838195	0.0662013	C2	0.580095	High	1	0	180	-62.2527
181	1.00 : 0.00 → High	High	Mote_181	0.57585	-0.366559	0.211824	-1.44988	0.55394	C3	0.537001	High	1	0	181	-56.2259
182	1.00 : 0.00 → High	High	Mote_182	1.29212	0.347553	0.884856	1.08644	-0.6898	C1	0.541447	High	1	0	182	-40.599
183	1.00 : 0.00 → High	High	Mote_183	0.0583812	0.117014	-1.53786	0.19033	0.292905	C3	0.551222	High	1	0	183	-55.2215
184	1.00 : 0.00 → High	High	Mote_184	-0.546544	0.997514	-1.5768	0.365056	0.446551	C2	0.518294	High	1	0	184	-60.1375
185	1.00 : 0.00 → High	High	Mote_185	0.0284407	-0.300432	0.123794	1.63042	-0.137672	C3	0.545678	High	1	0	185	-50.998
186	1.00 : 0.00 → High	High	Mote_186	-0.791928	1.72097	0.740884	-0.384426	0.307433	C1	0.521995	High	1	0	186	-62.7956
187	1.00 : 0.00 → High	High	Mote_187	-1.79049	-0.513046	1.02386	-0.727674	0.129238	C2	0.618984	High	1	0	187	-67.7241
188	0.00 : 1.00 → Medium	Medium	Mote_188	-2.05325	-1.01356	-0.0147578	-1.04036	-0.362815	C2	0.616487	Medium	0	1	188	-66.1336
189	1.00 : 0.00 → High	High	Mote_189	-2.03625	0.752204	-0.769755	-0.194565	-0.228337	C2	0.630776	High	1	0	189	-66.1586
190	1.00 : 0.00 → High	High	Mote_190	-1.50412	1.53664	1.7853	-0.063008	0.046575	C2	0.534841	High	1	0	190	-65.4429
191	0.00 : 1.00 → Medium	Medium	Mote_191	-2.60039	0.210821	0.201123	1.8452	-0.415023	C2	0.61283	Medium	0	1	191	-65.8352
192	1.00 : 0.00 → High	High	Mote_192	0.84701	-0.164662	0.921865	0.119927	0.266101	C3	0.499307	High	1	0	192	-50.7738
193	1.00 : 0.00 → High	High	Mote_193	-1.53908	-0.13237	-0.419612	-1.00117	-0.126499	C2	0.618293	High	1	0	193	-64.5436
194	1.00 : 0.00 → High	High	Mote_194	-0.707607	0.896983	0.49286	0.388067	0.123688	C2	0.509885	High	1	0	194	-59.63
195	1.00 : 0.00 → High	High	Mote_195	0.867039	-1.40446	-1.08931	0.658285	-0.619842	C3	0.616477	High	1	0	195	-43.0963
196	1.00 : 0.00 → High	High	Mote_196	0.337762	2.29853	-0.623419	-0.462944	0.348193	C1	0.608121	High	1	0	196	-55.7812
197	1.00 : 0.00 → High	High	Mote_197	1.22296	1.58685	1.97336	0.816961	0.509924	C1	0.606974	High	1	0	197	-49.9733
198	1.00 : 0.00 → High	High	Mote_198	-0.989966	0.289909	-0.745373	-1.12097	0.422388	C2	0.579135	High	1	0	198	-64.8304
199	1.00 : 0.00 → High	High	Mote_199	-1.39555	0.730363	1.7522	0.606197	0.155124	C2	0.556604	High	1	0	199	-64.2313
200	0.00 : 1.00 → Medium	Medium	Mote_200	-2.45762	1.61626	0.66201	-0.950229	-0.514431	C2	0.595199	Medium	0	1	200	-68.7739

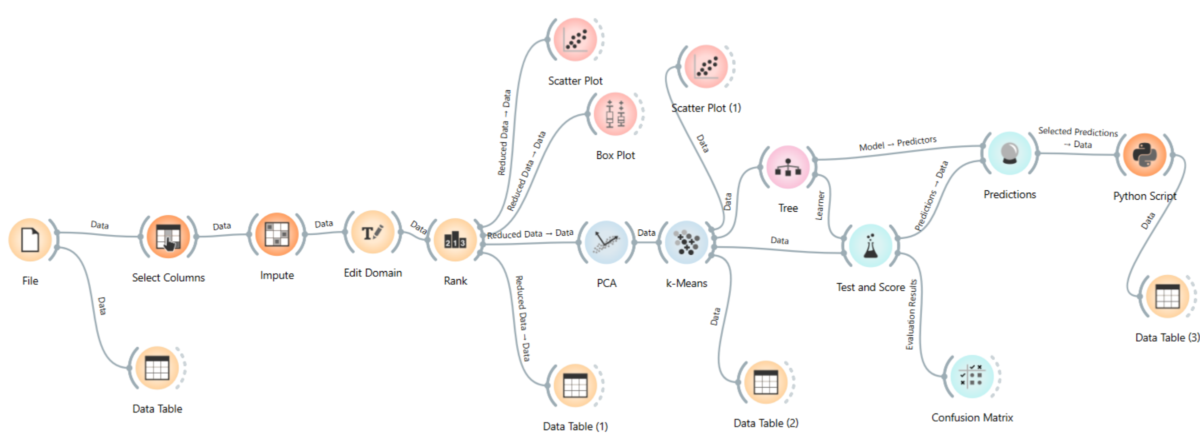
Show perfomance scores | Target class: (Average over classes)

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	1.000	1.000	1.000	1.000	1.000	1.000

Python Script:

	Interference_Level	Routing Decision	Signal_Strength (dBm)	Packet_Loss (%)	Y_Coordinate	Humidity (%)	Temperature (°C)
1	Medium	Enable Adaptive Frequency Hopping	-66.9063	34.945	6.42032	48.4468	29.1448
2	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-42.9234	55.419	0.8414	53.9295	18.0508
3	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-54.8424	53.887	1.61629	48.8505	26.5258
4	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-45.2063	57.3222	8.98554	44.4351	27.1343
5	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-60.3985	47.6671	6.06429	46.0318	23.4826
6	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-43.1343	56.5878	0.0919705	63.0389	29.7289
7	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-58.3239	48.5988	1.01472	50.3167	33.6873
8	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-69.6749	38.8171	6.63502	58.2135	33.5114
9	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-42.8385	52.4967	0.0506158	65.1709	24.0168
10	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-67.2614	37.6328	1.60808	74.5197	17.2648
11	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-60.4206	41.7915	5.48734	41.9656	34.6968
12	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-41.4981	59.8767	6.91895	79.9705	31.778
13	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-41.4818	59.4406	6.51961	71.395	17.4933
14	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-52.7969	47.5974	2.24269	44.088	33.4168
15	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-51.0449	56.0109	7.12179	83.6789	32.3979
16	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-56.5466	52.7058	2.37249	86.0436	25.3768
17	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-61.2037	40.6021	3.254	43.0539	26.8255
18	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-60.1401	45.5394	7.46491	53.8439	22.9801
19	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-49.8244	59.1549	6.49633	80.3101	16.0952
20	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-47.4288	50.3395	8.49223	77.413	21.7039
21	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-46.2526	56.9742	6.57613	49.2261	31.0571
22	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-46.3115	52.9735	5.68309	50.4675	15.0926
23	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-67.2638	41.9801	0.936748	58.5236	21.67
24	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-55.1674	54.5432	3.67716	64.2261	22.9634
25	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-68.2732	41.1694	2.65202	70.9127	25.7479
26	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-53.5141	51.228	2.4399	58.4457	33.3971
27	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-56.7541	51.8663	9.73011	63.1267	21.9269
28	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-43.3689	58.4455	3.93098	77.3735	21.9391
29	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-59.4725	43.7185	8.92047	41.8342	29.75
30	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-66.488	41.8012	6.31139	52.6218	24.0444
31	Medium	Enable Adaptive Frequency Hopping	-65.7102	34.6598	7.94811	75.6675	19.4921
32	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-47.1547	55.9627	5.02637	84.7603	24.0488
33	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-51.4535	50.8466	5.76904	65.5839	17.8171
34	Medium	Enable Adaptive Frequency Hopping	-66.9663	34.2393	4.92518	66.6057	18.5277
35	Medium	Enable Adaptive Frequency Hopping	-67.4768	33.2927	1.95243	45.3586	24.9674
36	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-48.9709	56.9629	7.22452	62.3706	23.3785
37	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-67.8171	35.5816	2.80772	66.6309	33.2969
38	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-45.3442	57.2477	0.24316	52.1235	22.2479
39	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-48.8127	50.6536	6.45472	53.4622	26.6118
40	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-67.5595	35.5934	1.77111	58.8642	27.6453
41	High	Switch to Low-Power Mode & Reduce Packet Forwarding	-67.4549	37.94	9.40459	41.0036	15.2619

Final Workflow



5. Expected Outcomes

- **40% Reduction in Packet Loss:** Improved signal transmission efficiency, ensuring fewer retransmissions. *(Example: In precision agriculture, soil moisture sensors experience reduced data loss, leading to more accurate irrigation scheduling.)*

- **AI-Optimized Routing:** Real-time adjustments based on network conditions, reducing congestion and improving efficiency. *(Example: In an environmental monitoring system, temperature and humidity data reach servers faster for real-time analysis.)*
- **Cluster-Based Topology Redesign:** Optimized routing paths based on sensor clustering, leading to better data flow. *(Example: Grouping sensors in dense networks ensures efficient data transmission, minimizing interference.)*
- **Scalability for IoT & Agriculture:** The system can be adapted for various IoT applications, such as automated irrigation and real-time pollution monitoring. *(Example: Smart Dust sensors in industrial IoT can dynamically adjust routes to ensure minimal data loss in factory environments.)*

6. Conclusion

This Orange-based workflow provides an efficient data-driven approach to minimizing packet loss, ensuring robust submillimeter sensor communication in Smart Dust Networks. By leveraging AI-driven techniques, including clustering, decision trees, and adaptive routing, the project achieves significant improvements in data reliability and network efficiency. The workflow is scalable and adaptable, making it suitable for a wide range of IoT applications, including precision agriculture and environmental monitoring.