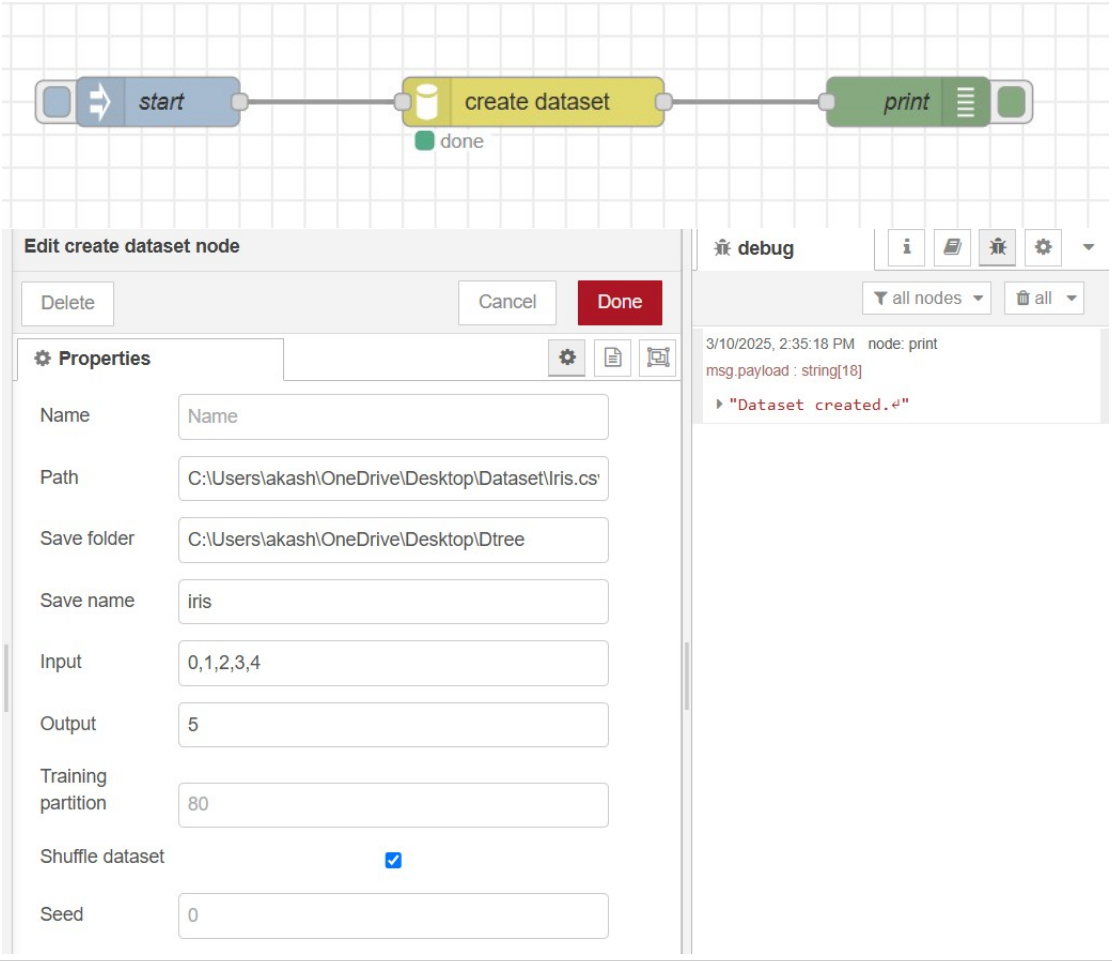
CHALLENGING TASK-4

T.Akash Reddy
21MIS1128

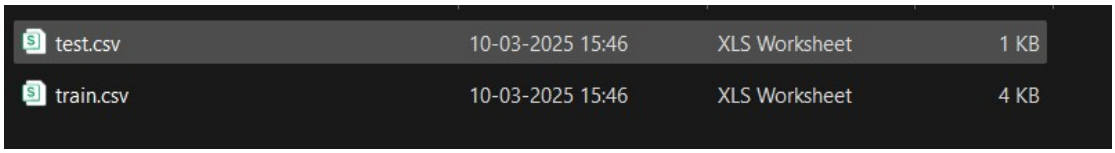
Decision tree classifier using node-red for iris dataset

Flow-1:

1. Here after installing the node-red machine learning modules, created a flow with an inject node, create dataset and a debug node.
2. Next in create dataset node we need to give the dataset path and all the necessary inputs and output column numbers.
3. Deploy and start the inject node will get to see a debug message that dataset has created.

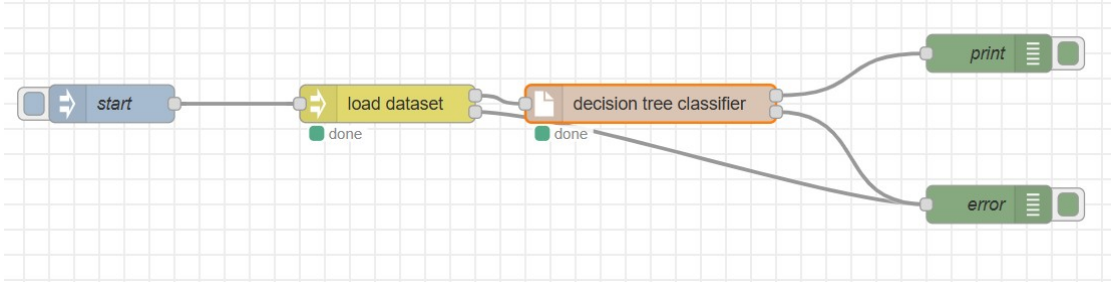


4. This flow creates a folder in which the test and train dataset will be stored.

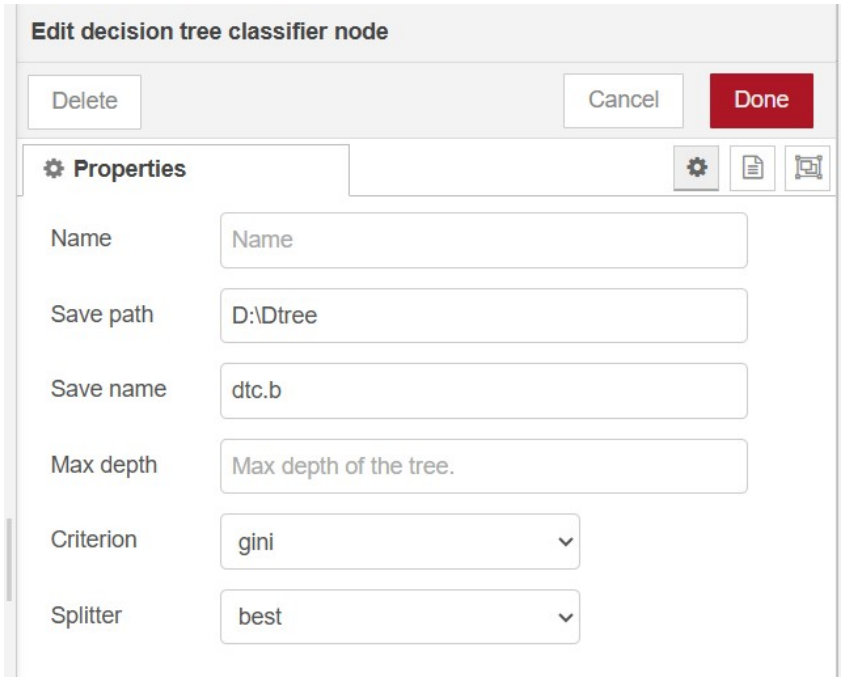


Flow-2:

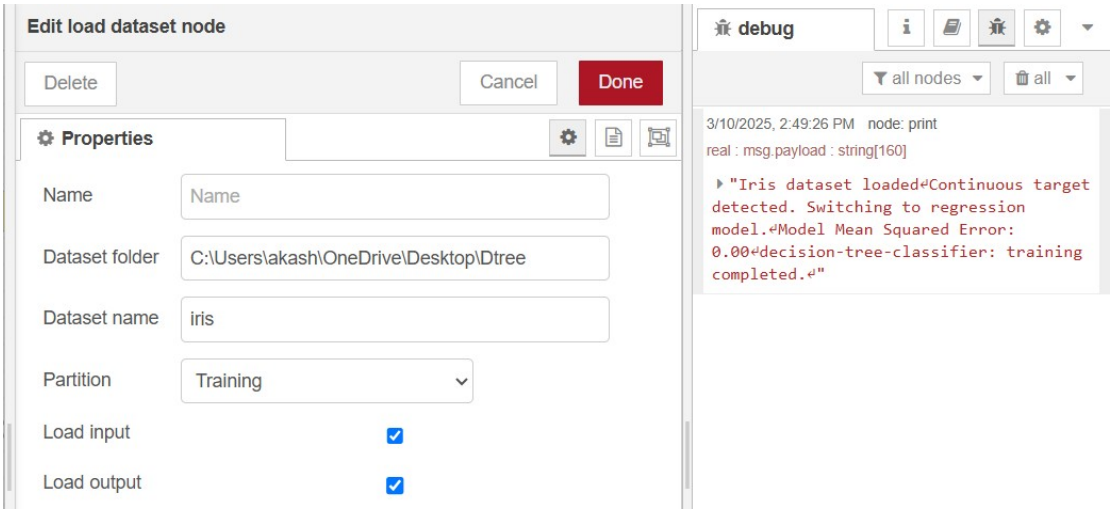
1. Here we need to add two new nodes which are load dataset and decision tree classifier.
2. In load dataset we need to set the path where the training and testing dataset had been stored.

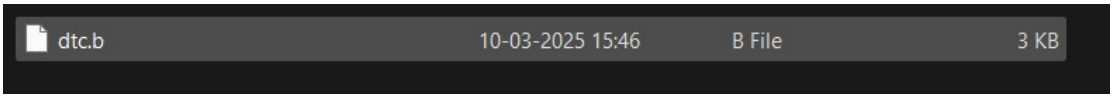


3. In decision tree node we need to give a path where trained model will be stored.



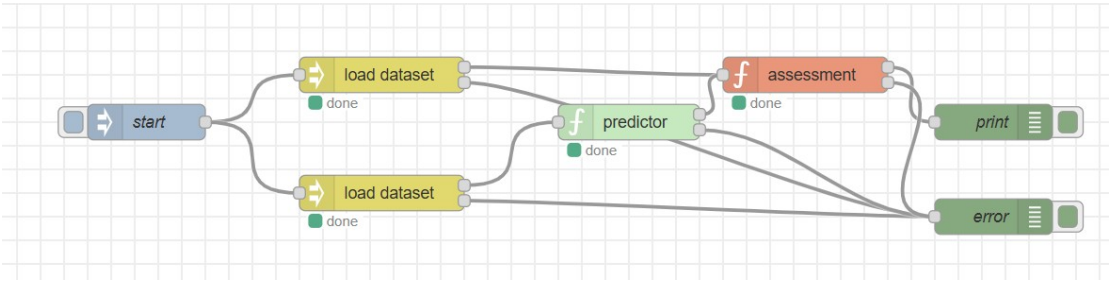
4. On successful deploy we can see a message in debug window that model has been trained and a dtc.b file will be stored.

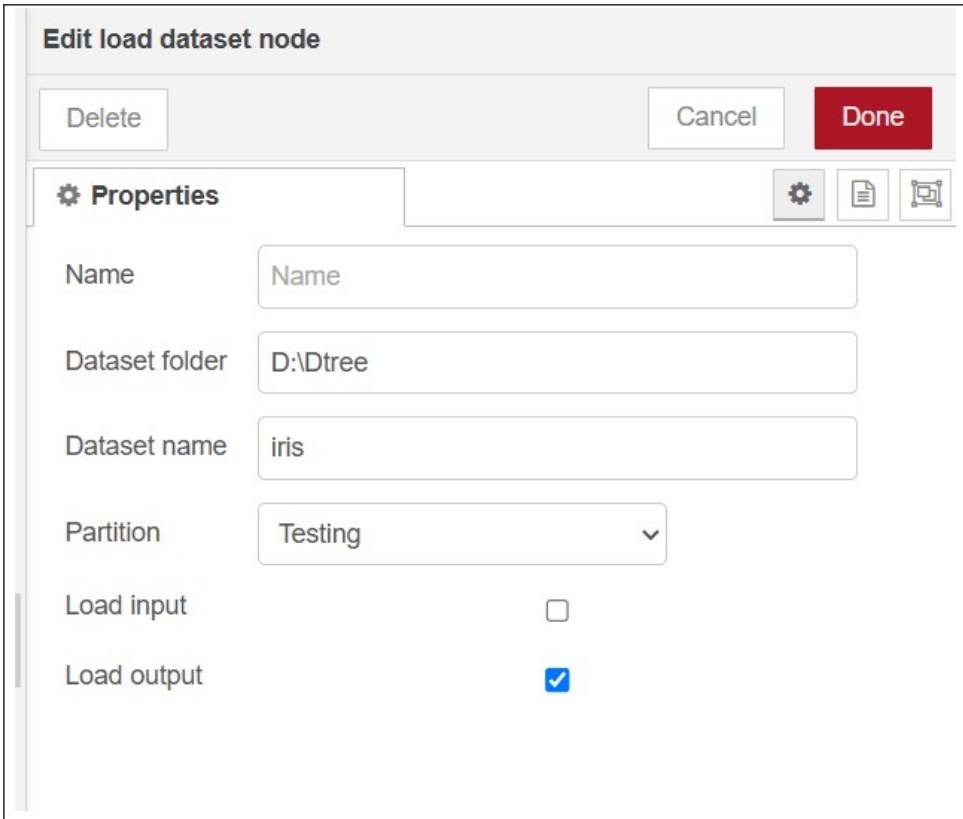


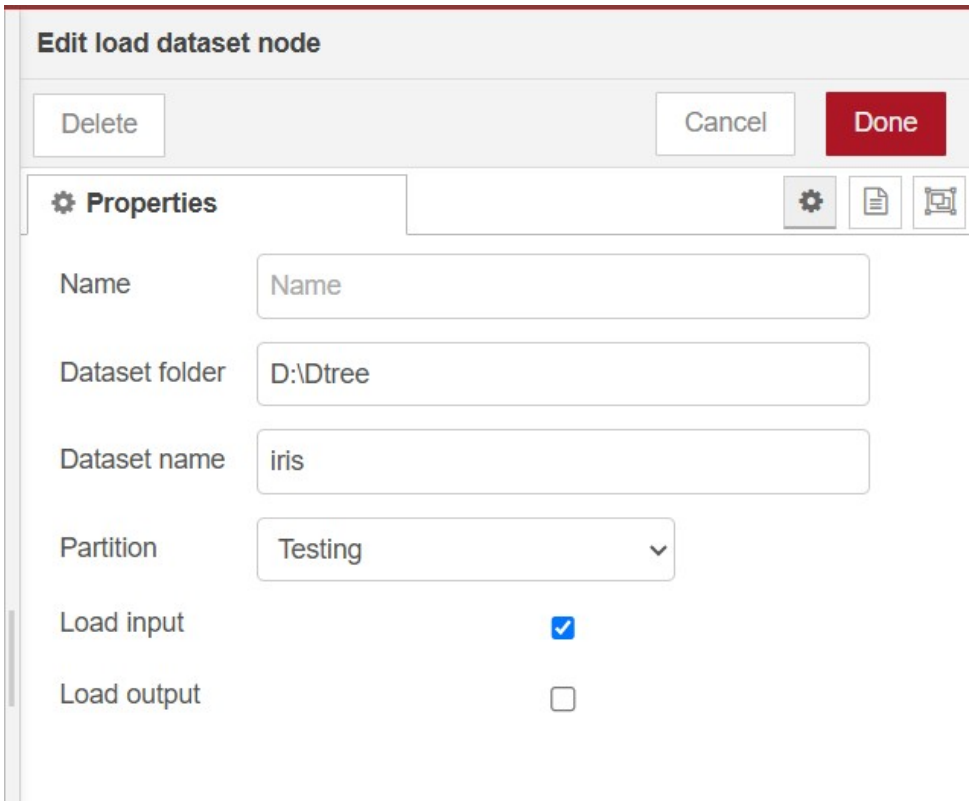


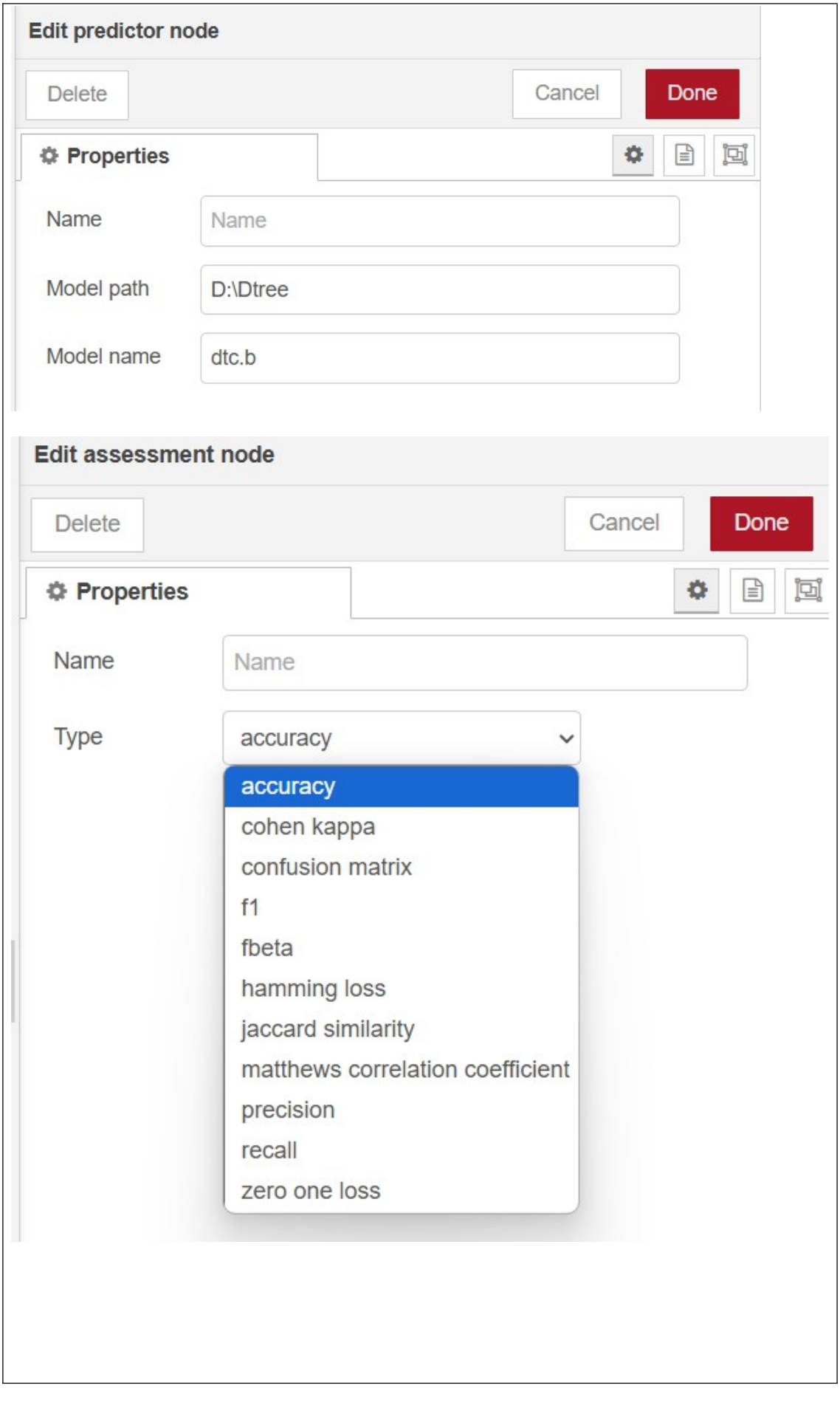
Flow-3:

1. Here we need to add two load dataset nodes to pass the training and testing dataset to the predictor node which takes the model trained in flow 2 and gives the accuracy of the model in debug window.
2. Assessment node helps us in selecting different variety of score checker like accuracy, cohen kappa, f1 score and etc..
3. Below are the screen shots attached for better understanding.




Edit load dataset node
Delete
Cancel
Done
Properties
Name
Name
Dataset folder
D:\Dtree
Dataset name
iris
Partition
Testing
Load input
Load output


Edit load dataset node
Delete
Cancel
Done
Properties
Name
Name
Dataset folder
D:\Dtree
Dataset name
iris
Partition
Testing
Load input
Load output

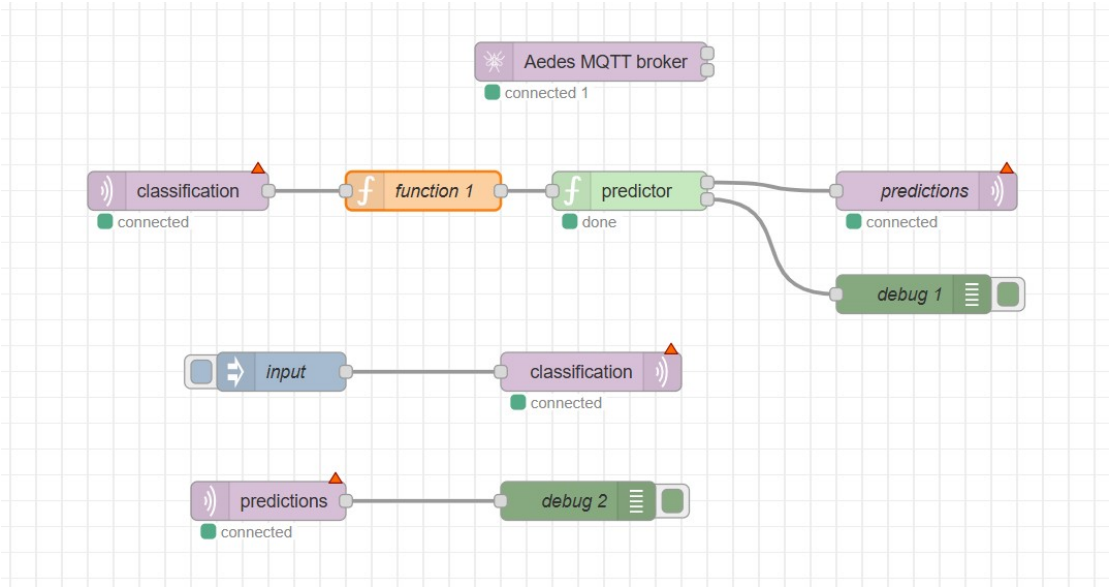

Edit predictor node
Delete
Cancel
Done
Properties
Name
Name
Model path
D:\Dtree
Model name
dtc.b
Edit assessment node
Delete
Cancel
Done
Properties
Name
Name
Type
accuracy
accuracy
cohen kappa
confusion matrix
f1
fbeta
hamming loss
jaccard similarity
matthews correlation coefficient
precision
recall
zero one loss

```
3/10/2025, 3:48:25 PM   node: print
predicted : msg.payload : number
0.9333333333333333
```

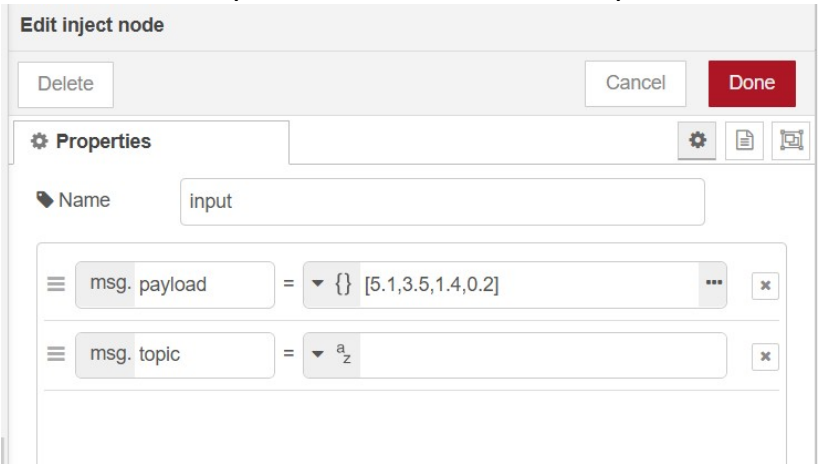
We can see the accuracy of the model is 94% approx.

Flow-4:

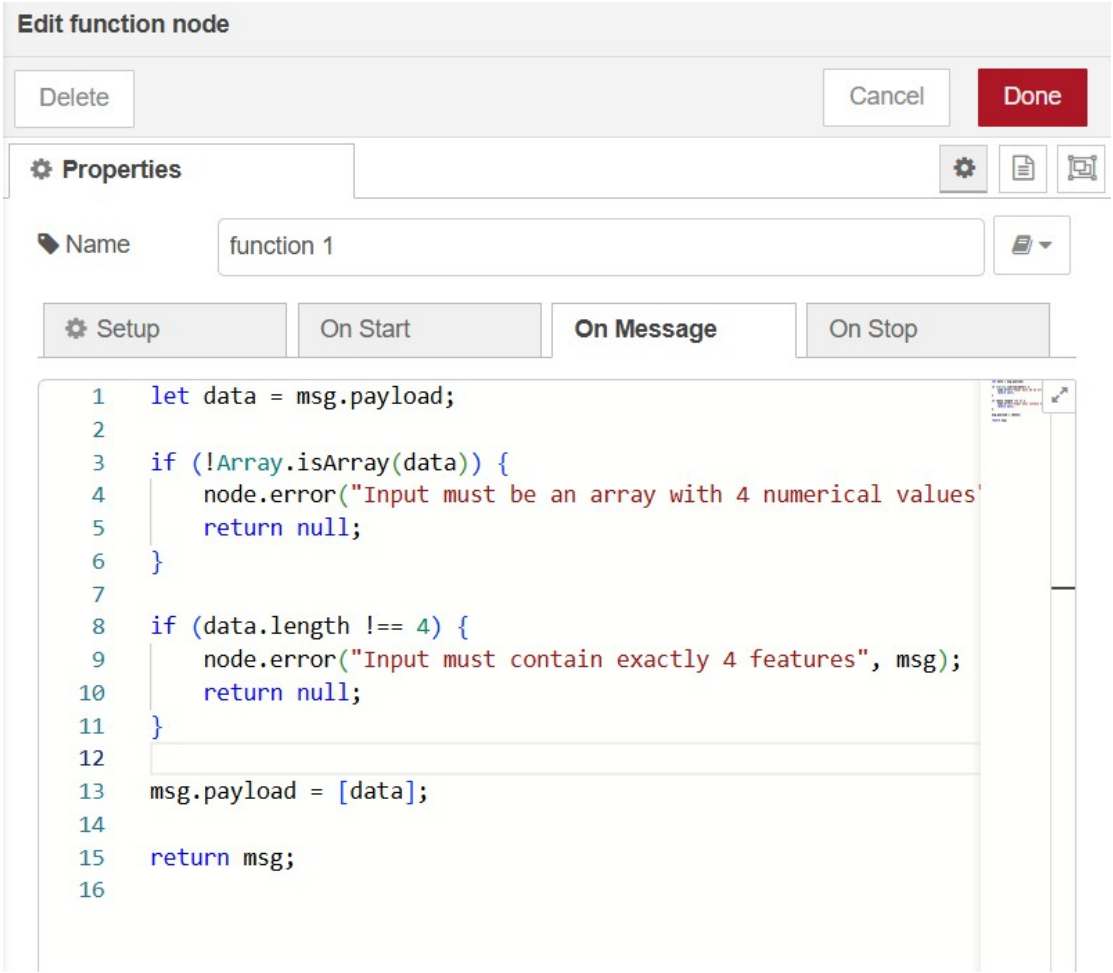
1. This flow is used to check the prediction of the model using the user input via mqtt borker.



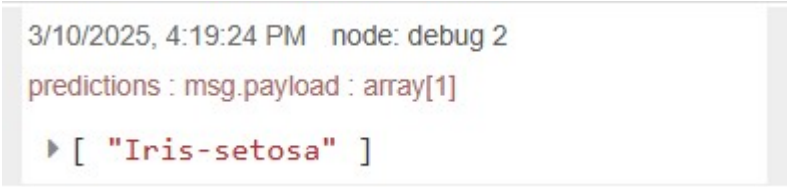
2. In input node I had passed the feature values in a json 1d array format and then the input node connected to a publisher.



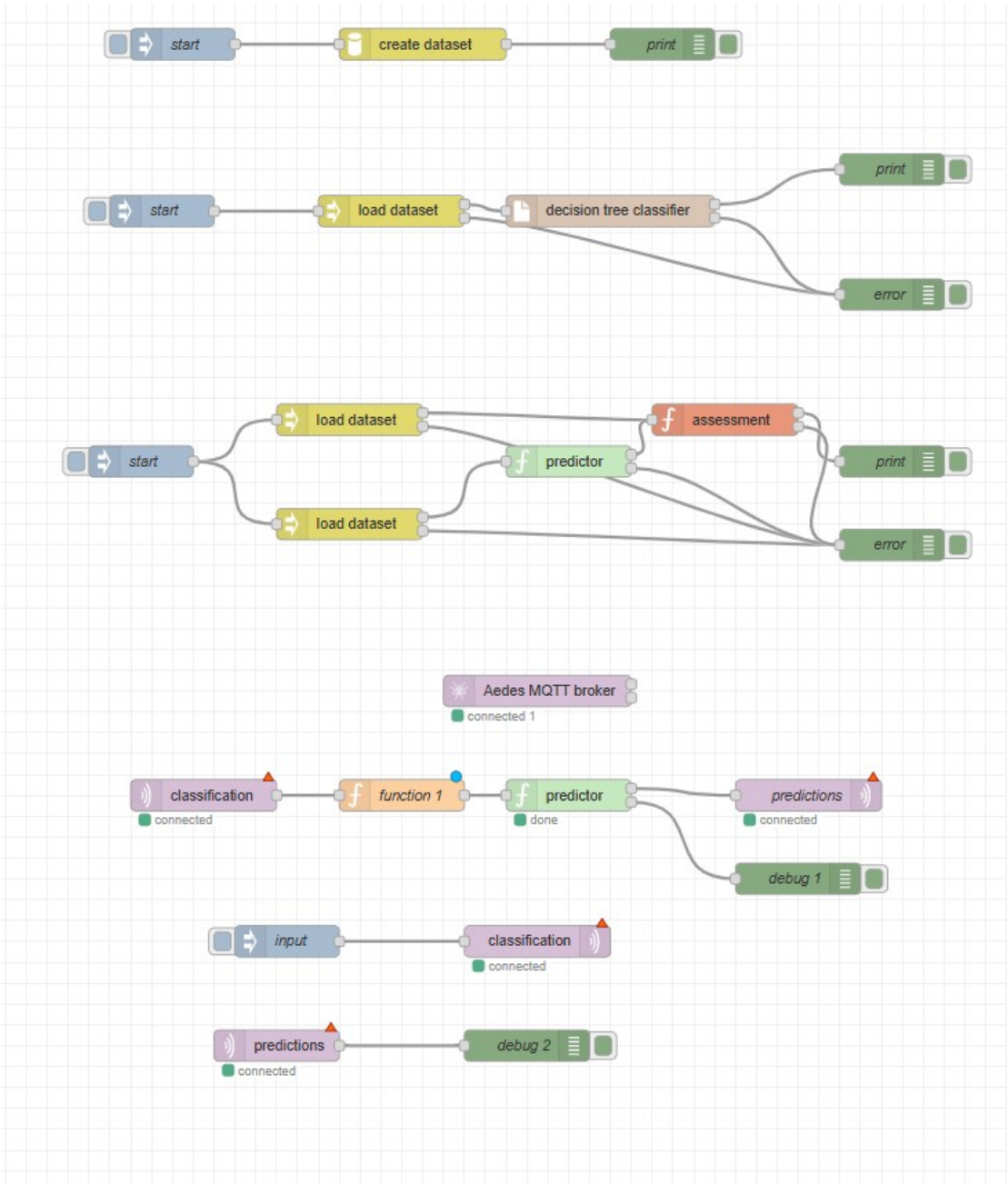
3. Then the subscriber subscribes the input from the publisher and pass it to a function node for checking data validation and to convert the 1d array input to 2d array because the skilearn expects the input in 2d array to pass to the decision tree model.



4. Then the predictor predicts the output and debugs the output via publisher and a subscriber and below is the output that we got for the given input.



Overall node-red flow:



Note: Initially the paths I used C/users/akash/Onedrive/Desktop then I changed the paths to D drive because of failed permissions due to onedrive.

Json code for the above flow:


```
[
  {
    "id": "873554034edb1d14",
    "type": "tab",
    "label": "Flow 1",
    "disabled": false,
    "info": "",
    "env": []
  },
  {
    "id": "1b92c9be813bb468",
    "type": "inject",
    "z": "873554034edb1d14",
    "name": "start",
    "props": [
      {
        "p": "payload"
      },
      {
        "p": "topic",
        "vt": "str"
      }
    ],
    "repeat": "",
    "crontab": "",
    "once": false,
    "onceDelay": 0.1,
    "topic": "",
    "payload": "",
    "payloadType": "date",
    "x": 230,
    "y": 120,
    "wires": [
      [
        "b6ef1785e0021b5f"
      ]
    ]
  },
  {
    "id": "b6ef1785e0021b5f",
    "type": "create dataset",
    "z": "873554034edb1d14",
    "name": "",
    "path": "D:\\Dataset\\Iris.csv",
    "saveFolder": "D:\\Dtree",
    "saveName": "iris",
    "input": "0,1,2,3",
    "output": "4",
    "trainingPartition": "",
    "shuffle": true,
```

```
    "seed": "",
    "x": 460,
    "y": 120,
    "wires": [
      [
        "cd42b17f44d363f4"
      ]
    ]
  },
  {
    "id": "cd42b17f44d363f4",
    "type": "debug",
    "z": "873554034edb1d14",
    "name": "print",
    "active": true,
    "tosidebar": true,
    "console": false,
    "tostatus": false,
    "complete": "payload",
    "targetType": "msg",
    "statusVal": "",
    "statusType": "auto",
    "x": 690,
    "y": 120,
    "wires": []
  },
  {
    "id": "108356eeeff8cf48",
    "type": "inject",
    "z": "873554034edb1d14",
    "name": "start",
    "props": [
      {
        "p": "payload"
      },
      {
        "p": "topic",
        "vt": "str"
      }
    ],
    "repeat": "",
    "crontab": "",
    "once": false,
    "onceDelay": 0.1,
    "topic": "",
    "payload": "",
    "payloadType": "date",
    "x": 210,
    "y": 280,
    "wires": [
      [
        "405f10ed799d42f1"
      ]
    ]
  },
  {
    "id": "405f10ed799d42f1",
```

```
    "type": "load dataset",
    "z": "873554034edb1d14",
    "name": "",
    "datasetFolder": "D:\\Dtree",
    "datasetName": "iris",
    "partition": "train.csv",
    "input": true,
    "output": true,
    "x": 430,
    "y": 280,
    "wires": [
      [
        "a09f27bc85747f0d"
      ],
      [
        "9ac1946936af6f30"
      ]
    ]
  },
  {
    "id": "a09f27bc85747f0d",
    "type": "decision tree classifier",
    "z": "873554034edb1d14",
    "name": "",
    "savePath": "D:\\Dtree",
    "saveName": "dtc.b",
    "maxDepth": "",
    "criterion": "gini",
    "splitter": "best",
    "x": 640,
    "y": 280,
    "wires": [
      [
        "bdd106046a4f4bde"
      ],
      [
        "9ac1946936af6f30"
      ]
    ]
  },
  {
    "id": "bdd106046a4f4bde",
    "type": "debug",
    "z": "873554034edb1d14",
    "name": "print",
    "active": true,
    "tosidebar": true,
    "console": false,
    "tostatus": false,
    "complete": "payload",
    "targetType": "msg",
    "statusVal": "",
    "statusType": "auto",
    "x": 910,
    "y": 240,
    "wires": []
  },
```

```
{
    "id": "9ac1946936af6f30",
    "type": "debug",
    "z": "873554034edb1d14",
    "name": "error",
    "active": true,
    "tosidebar": true,
    "console": false,
    "tostatus": false,
    "complete": "payload",
    "targetType": "msg",
    "statusVal": "",
    "statusType": "auto",
    "x": 910,
    "y": 360,
    "wires": []
},
{
    "id": "6ec7228f17aca967",
    "type": "inject",
    "z": "873554034edb1d14",
    "name": "start",
    "props": [
        {
            "p": "payload"
        },
        {
            "p": "topic",
            "vt": "str"
        }
    ],
    "repeat": "",
    "crontab": "",
    "once": false,
    "onceDelay": 0.1,
    "topic": "",
    "payload": "",
    "payloadType": "date",
    "x": 190,
    "y": 520,
    "wires": [
        [
            "a4dd923d5fff0418",
            "4a5f05cb271445d9"
        ]
    ]
},
{
    "id": "a4dd923d5fff0418",
    "type": "load dataset",
    "z": "873554034edb1d14",
    "name": "",
    "datasetFolder": "D:\\Dtree",
    "datasetName": "iris",
    "partition": "test.csv",
    "input": false,
    "output": true,
```

```
    "x": 390,
    "y": 480,
    "wires": [
      [
        "8780ce9fa71ecd0d"
      ],
      [
        "3fa762ec7ba7d076"
      ]
    ]
  },
  {
    "id": "4a5f05cb271445d9",
    "type": "load dataset",
    "z": "873554034edb1d14",
    "name": "",
    "datasetFolder": "D:\\Dtree",
    "datasetName": "iris",
    "partition": "test.csv",
    "input": true,
    "output": false,
    "x": 390,
    "y": 580,
    "wires": [
      [
        "f8230cc9dd79c090"
      ],
      [
        "3fa762ec7ba7d076"
      ]
    ]
  },
  {
    "id": "f8230cc9dd79c090",
    "type": "predictor",
    "z": "873554034edb1d14",
    "name": "",
    "modelPath": "D:\\Dtree",
    "modelName": "dtc.b",
    "x": 600,
    "y": 520,
    "wires": [
      [
        "8780ce9fa71ecd0d"
      ],
      [
        "3fa762ec7ba7d076"
      ]
    ]
  },
  {
    "id": "8780ce9fa71ecd0d",
    "type": "assessment",
    "z": "873554034edb1d14",
    "name": "",
    "score": "accuracy_score",
    "x": 750,
```

```
    "y": 480,
    "wires": [
      [
        "6440f48c61b71e6c"
      ],
      [
        "3fa762ec7ba7d076"
      ]
    ]
  },
  {
    "id": "6440f48c61b71e6c",
    "type": "debug",
    "z": "873554034edb1d14",
    "name": "print",
    "active": true,
    "tosidebar": true,
    "console": false,
    "tostatus": false,
    "complete": "payload",
    "targetType": "msg",
    "statusVal": "",
    "statusType": "auto",
    "x": 910,
    "y": 520,
    "wires": []
  },
  {
    "id": "3fa762ec7ba7d076",
    "type": "debug",
    "z": "873554034edb1d14",
    "name": "error",
    "active": true,
    "tosidebar": true,
    "console": false,
    "tostatus": false,
    "complete": "payload",
    "targetType": "msg",
    "statusVal": "",
    "statusType": "auto",
    "x": 910,
    "y": 600,
    "wires": []
  },
  {
    "id": "e1d5ee9107497276",
    "type": "mqtt in",
    "z": "873554034edb1d14",
    "name": "",
    "topic": "classification",
    "qos": "2",
    "datatype": "auto-detect",
    "broker": "56469696e18baa31",
    "nl": false,
    "rap": true,
    "rh": 0,
    "inputs": 0,
```

```
    "x": 250,
    "y": 840,
    "wires": [
      [
        "2ac761f04df8ca8e"
      ]
    ]
  },
  {
    "id": "6fc47e82de6b5cee",
    "type": "mqtt out",
    "z": "873554034edb1d14",
    "name": "predictions",
    "topic": "predictions",
    "qos": "",
    "retain": "",
    "respTopic": "",
    "contentType": "",
    "userProps": "",
    "correl": "",
    "expiry": "",
    "broker": "56469696e18baa31",
    "x": 830,
    "y": 840,
    "wires": []
  },
  {
    "id": "f8063de86208d1cc",
    "type": "predictor",
    "z": "873554034edb1d14",
    "name": "",
    "modelPath": "D:\\Dtree",
    "modelName": "dtc.b",
    "x": 600,
    "y": 840,
    "wires": [
      [
        "6fc47e82de6b5cee"
      ],
      [
        "7f3190f7fe10e7cb"
      ]
    ]
  },
  {
    "id": "7f3190f7fe10e7cb",
    "type": "debug",
    "z": "873554034edb1d14",
    "name": "debug 1",
    "active": true,
    "tosidebar": true,
    "console": false,
    "tostatus": false,
    "complete": "false",
    "statusVal": "",
    "statusType": "auto",
    "x": 820,
```

```
    "y": 920,
    "wires": []
  },
  {
    "id": "a11e7f5354c616f7",
    "type": "inject",
    "z": "873554034edb1d14",
    "name": "input",
    "props": [
      {
        "p": "payload"
      },
      {
        "p": "topic",
        "vt": "str"
      }
    ],
    "repeat": "",
    "crontab": "",
    "once": false,
    "onceDelay": 0.1,
    "topic": "",
    "payload": "[5.1,3.5,1.4,0.2]",
    "payloadType": "json",
    "x": 330,
    "y": 980,
    "wires": [
      [
        "14769e956fdb309f"
      ]
    ]
  },
  {
    "id": "106762c64fdedfba",
    "type": "aedes broker",
    "z": "873554034edb1d14",
    "name": "",
    "mqtt_port": 1883,
    "mqtt_ws_bind": "port",
    "mqtt_ws_port": "",
    "mqtt_ws_path": "",
    "cert": "",
    "key": "",
    "certname": "",
    "keyname": "",
    "dburl": "",
    "usetls": false,
    "x": 570,
    "y": 740,
    "wires": [
      [],
      []
    ]
  },
  {
    "id": "14769e956fdb309f",
    "type": "mqtt out",
```

```
    "z": "873554034edb1d14",
    "name": "",
    "topic": "classification",
    "qos": "",
    "retain": "",
    "respTopic": "",
    "contentType": "",
    "userProps": "",
    "correl": "",
    "expiry": "",
    "broker": "56469696e18baa31",
    "x": 570,
    "y": 980,
    "wires": []
  },
  {
    "id": "d9a7e833058988d3",
    "type": "mqtt in",
    "z": "873554034edb1d14",
    "name": "",
    "topic": "predictions",
    "qos": "2",
    "datatype": "auto-detect",
    "broker": "56469696e18baa31",
    "nl": false,
    "rap": true,
    "rh": 0,
    "inputs": 0,
    "x": 320,
    "y": 1080,
    "wires": [
      [
        "ad3b5e6a7c70cca5"
      ]
    ]
  },
  {
    "id": "ad3b5e6a7c70cca5",
    "type": "debug",
    "z": "873554034edb1d14",
    "name": "debug 2",
    "active": true,
    "tosidebar": true,
    "console": false,
    "tostatus": false,
    "complete": "payload",
    "targetType": "msg",
    "statusVal": "",
    "statusType": "auto",
    "x": 560,
    "y": 1080,
    "wires": []
  },
  {
    "id": "2ac761f04df8ca8e",
    "type": "function",
    "z": "873554034edb1d14",
```

```
    "name": "function 1",
    "func": "let data = msg.payload;\n\nif (!Array.isArray(data)) {\n    node.error(\"Input must be an array with 4 numerical values\", msg);\n    return null;\n}\n\nif (data.length !== 4) {\n    node.error(\"Input must contain exactly 4 features\", msg);\n    return null;\n}\n\nmsg.payload = [data];\n\nreturn msg;\n",
    "outputs": 1,
    "timeout": 0,
    "noerr": 0,
    "initialize": "",
    "finalize": "",
    "libs": [],
    "x": 440,
    "y": 840,
    "wires": [
      [
        "f8063de86208d1cc"
      ]
    ]
  },
  {
    "id": "56469696e18baa31",
    "type": "mqtt-broker",
    "name": "",
    "broker": "",
    "port": 1883,
    "clientid": "",
    "autoConnect": true,
    "usetls": false,
    "protocolVersion": 4,
    "keepalive": 60,
    "cleansession": true,
    "autoUnsubscribe": true,
    "birthTopic": "",
    "birthQos": "0",
    "birthRetain": "false",
    "birthPayload": "",
    "birthMsg": {},
    "closeTopic": "",
    "closeQos": "0",
    "closeRetain": "false",
    "closePayload": "",
    "closeMsg": {},
    "willTopic": "",
    "willQos": "0",
    "willRetain": "false",
    "willPayload": "",
    "willMsg": {},
    "userProps": "",
    "sessionExpiry": ""
  }
]
```