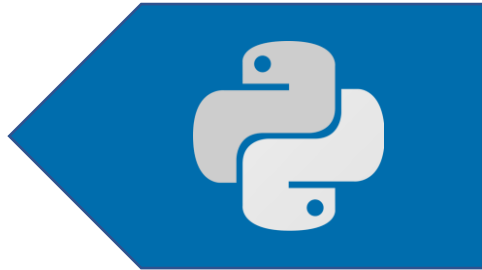
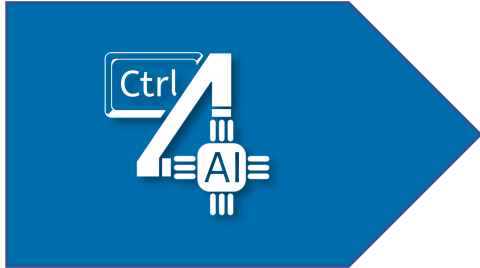


**A helper package for Machine Learning and Deep Learning solutions**



**Developers:** Shaji, Charu, Selva

---

<b>Package Name</b>	Ctrl4AI
<b>Latest Version</b>	1.0.2
<b>Environment</b>	Python
<b>Hosted On</b>	PyPi
<b>Mode of install</b>	pip

---



# What's new?

**Open Source Package** with emphasis on data preprocessing so far.

**Self intelligent methods** that can be employed at the levels of abstraction or customization.

The flow of **auto-preprocessing** is orchestrated compatible to the learning type.

**Parameter tuning** allows users to transform the data precisely to their specifications.

Developed **computations** to discover the type of data, it's distribution, correlation etc. which are handled in the background.

## ctrl4ai.automl

```
preprocess(dataset, learning_type, target_variable=None, target_type=None,
impute_null_method='central_tendency', tranform_categorical='label_encoding',
categorical_threshold=0.3, remove_outliers=False, log_transform=None,
drop_null_dominated=True, dropna_threshold=0.7, derive_from_datetime=True,
ohe_ignore_cols=[], feature_selection=True, define_continuous_cols=[],
define_categorical_cols=[])
```

```
dataset1_labels,dataset1_processed=automl.preprocess(dataset1,'supervised',target_variable='fare_amount',target_type='continuous')
dataset1_processed.head()
```

Dropping single valued column(s) new\_user,year\_of\_pickup\_datetime,year\_of\_dropoff\_datetime

Columns identified as continuous are tip\_amount,pickup\_longitude,pickup\_latitude,dropoff\_longitude,dropoff\_latitude

Columns identified as categorical are vendor\_id,tolls\_amount,mta\_tax,passenger\_count,rate\_code,store\_and\_fwd\_flag,payment\_type,surcharge,hour\_of\_pickup\_datetime,weekday\_of\_pickup\_datetime,month\_of\_pickup\_datetime

Replaced nulls in tolls\_amount with mode

Replaced nulls in mta\_tax with mode

Replaced nulls in passenger\_count with mode

Replaced nulls in rate\_code with mode

Replaced nulls in store\_and\_fwd\_flag with mode

Replaced nulls in payment\_type with mode

Replaced nulls in surcharge with mode

Labels for vendor\_id: {'DST000401': 0, 'DST000532': 1}

Labels for store\_and\_fwd\_flag: {'N': 0, 'Y': 1}

Labels for payment\_type: {'CRD': 0, 'CSH': 1, 'DIS': 2, 'NOC': 3, 'UNK': 4}

Replaced nulls in tip\_amount with mode

Replaced nulls in pickup\_longitude with mean

Replaced nulls in pickup\_latitude with mean

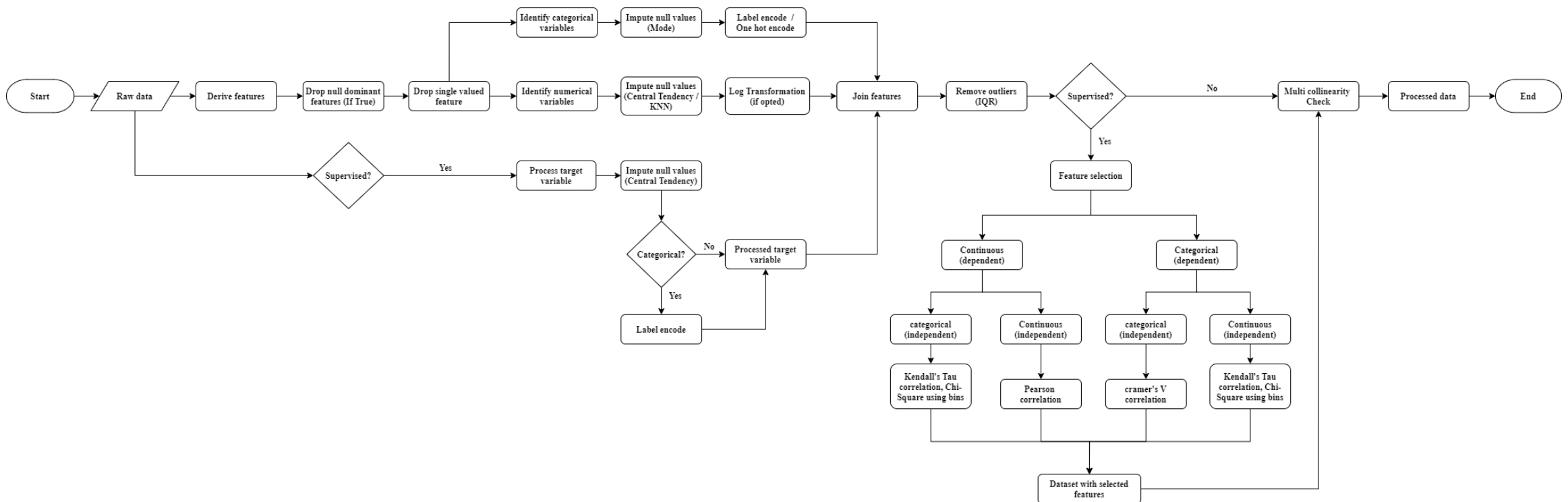
Replaced nulls in dropoff\_longitude with mean

Replaced nulls in dropoff\_latitude with mean

	pickup_longitude	dropoff_longitude	vendor_id	tolls_amount	passenger_count	rate_code	store_and_fwd_flag	surcharge	hour_of_pickup_datetime	weekday_of_pickup_datetime	month_of_pickup_datetime
0	-74.003939	-73.993369	0	0.0	1.0	1.0	0	0.5	4		0
1	-73.973864	-73.958701	0	0.0	3.0	1.0	0	0.0	18		6
2	-73.954406	-73.970780	0	0.0	2.0	1.0	0	0.0	8		0
3	-73.962345	-73.975512	1	0.0	2.0	1.0	0	0.0	9		4
4	-74.004657	-73.999369	0	0.0	1.0	1.0	0	0.0	13		2

# ctrl4ai 1.0.2

`pip install ctrl4ai`



**Don't repeat! Just Automate!**