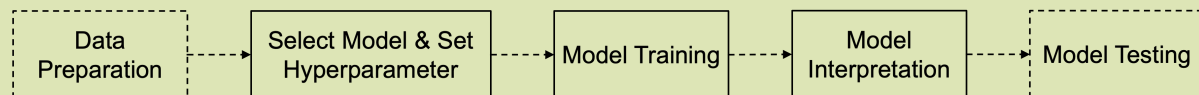


Interpretable Model Development :: CHEATSHEET

Modeva provides built-in algorithms for **developing inherently interpretable models**, spanning from linear regression and tree-based methods to neural networks and advanced mixture-of-experts models. User can set hyperparameters for controlling model complexity, effect sparsity, and feature monotonicity.



Refer to separate cheatsheets for **data preparation**, **hyperparameter tuning** and **model testing**.

Classification Models

- MoElasticNet()
- MoReLUENNClassifier()
- MoGAMINetClassifier()
- MoDecisionTreeClassifier()
- MoNeuralTreeClassifier()
- MoGLMTreeClassifier()
- MoGLMTreeBoostClassifier()
- MoRandomForestClassifier()

- MoGradientBoostingClassifier()
- MoXGBClassifier()
- MoLGBMClassifier()
- MoCatBoostClassifier()
- MoMoEClassifier()

Regression Models

- MoLogisticRegression()
- MoReLUENNRegressor()
- MoGAMINetRegressor()

- MoDecisionTreeRegressor()
- MoNeuralTreeRegressor()
- MoGLMTreeRegressor()
- MoGLMTreeBoostRegressor()
- MoRandomForestRegressor()
- MoGradientBoostingRegressor()
- MoXGBRegressor()
- MoLGBMRegressor()
- MoCatBoostRegressor()
- MoMoERegressor()

Built-in Interpretable Models:

```
from modeva.models import MoXXXXX
```

TestSuite Class for model interpretation

```
from modeva import TestSuite  
ts = TestSuite(ds, model)
```

Feature importance

```
ts.interpret_fi()
```

Global main and interaction effects

```
ts.interpret_effects()
```

Local feature importance

```
ts.interpret_local_fi()
```

Model-specific Interpretation

Generalized Linear Models (MoElasticNet, MoLogisticRegression): interpret linear coefficients

```
ts.interpret_coef() #Global  
ts.interpret_local_linear_fi() #Local
```

Decision Tree(MoDecisionTreeClassifier, MoDecisionTreeRegressor): tree visualization

```
ts.interpret_global_tree()  
ts.interpret_local_tree()
```

GAMI-Net (MoGAMINetClassifier, MoGAMINetRegressor) and **Tree Ensemble** (MoLGBMClassifier, MoLGBMRegressor, MoXGBClassifier, MoXGBRegressor, MoCatBoostClassifier, MoCatBoostRegressor, MoGradientBoostingClassifier, MoGradientBoostingRegressor, MoRandomForestClassifier, MoRandomForestRegressor): interpret FANOVA effect importance

```
ts.interpret_ei() # Global  
ts.interpret_local_ei() #Local
```

Model-specific Interpretation

ReLU-DNN (MoReLUENNClassifier, MoReLUENNRegressor): interpret local linear models

```
ts.interpret_llm_summary()  
ts.interpret_llm_pc() #LLM parallel coordinate plot  
ts.interpret_llm_profile() #LLM profile plot  
ts.interpret_local_linear_fi() #Local FI per sample
```

MoE (Mixture of Experts) (MoMoEClassifier, MoMoERegressor): clusterwise interpretation

```
ts.interpret_local_moe_weights()  
ts.interpret_moe_cluster_analysis()  
res = ts.interpret_fi()  
res = ts.interpret_ei()  
res = ts.interpret_effects()  
res = ts.interpret_local_fi()  
res = ts.interpret_local_ei()  
res.plot("k") # cluster-k visualization
```