

Squark Seer Leaderboard Explained

Squark Seer produces a Leaderboard that lists the best-performing models that were trained on your specific data from Squark's set of powerful codeless AI algorithms. While Squark Seer may have built thousands of models while you waited for results, the Leaderboard only contains the most accurate model for each algorithm we used. For example, if a Deep Learner displays in the leaderboard, then it is the most accurate Deep Learning model we created, out of perhaps thousands of deep learners built for your data. Since Squark's Leaderboard only contains the most accurate instances of models and their underlying algorithms, if a model algorithm is absent it is because it was not used.

Squark cross-validates more than 15 algorithms that are applied to your data automatically. These algorithms include fixed, specific, and dynamic grids and multiple instances of algorithms including: XGBoost Gradient Boosting Machines, other Gradient Boosting Machines, general linear models (GLMs), multiple "Tree" methods such as Distributed Random Forests, Extreme Trees, & Isolation Trees, multiple Deep Neural Networks, and multiple types of Ensemble Models.

Each model/algorithm is listed in order of accuracy using a default metric. Squark uses the metric "Area Under the Curve" (AUC) for binary classification, the metric "Mean per Class Error" for multiclass classification, and the metric "Residual Deviance" for Regression.

Squark provides a number of other metrics, which may be helpful for advanced users and data scientists including:

- **Area Under the Curve or AUC** (in Binary Classification Only) is used to evaluate how well a binary classification model is able to distinguish between true positives and false positives. An AUC of 1 indicates a perfect classifier, while an AUC of .5 indicates a poor classifier, whose performance is no better than random guessing.
- **Mean Per Class Error** (in Multi-class Classification only) is the average of the errors of each class in your multiclass dataset. This metric speaks toward misclassification of the data across the classes. The lower this metric, the better.
- **Residual Deviance** (in Regression Only) is short for Mean Residual Deviance and measures the goodness of the model's fit. In a perfect world, this metric would be zero. Deviance is equal to MSE in Gaussian distributions. If Deviance doesn't equal MSE, then it gives a more useful estimate of error, which is why Squark uses it as the default metric to rank for regression models.
- **Confusion Matrix**, if calculated, is a table depicting performance of the model used for predictions in the context of the false positives, false negatives, true positives, and true negatives, generated via cross-validation
- **Logloss** (or Logarithmic Loss) measures classification performance, specifically, uncertainty. This metric evaluates how closely a model's predicted values are to the actual target value. For example, does a model tend to assign a high predicted value like .90 for the positive class, or does it show a poor ability to identify the positive class and assign a lower predicted value like .40? Logloss ranges between 0 and 1, with 0 meaning that the model correctly assigns a probability of 0% or 100%. Logloss is sensitive to low probabilities that are erroneous.
- **MAE** or the Mean Absolute Error is an average of the absolute errors. The smaller the MAE the better the model's performance. The MAE units are the same units as your data's dependent variable/target (so if that's dollars, this is in dollars), which is useful for understanding whether the size of the error is meaningful or not. MAE is not sensitive to outliers. If your data has a lot of outliers, then examine the root mean squared error (RMSE), which is sensitive to outliers.
- **MSE** is the Mean Square Error and is a model quality metric. Closer to zero is better. The MSE metric measures the average of the squares of the errors or deviations. MSE takes the distances from the points to the regression line (these distances are the "errors") and then squares them to remove any negative signs. MSE incorporates both the variance and the bias of the predictor. MSE gives more weight to larger differences in errors than MAE.
- **RMSE** is the Root Mean Square Error. The RMSE will always be larger or equal to the MAE. The RMSE metric evaluates how well a model can predict a continuous value. The RMSE units are the same units as your data's dependent variable/target (so if that's dollars, this is in dollars), which is useful for understanding whether the size of the error is meaningful or not. The smaller the RMSE, the better the model's performance. RMSE is sensitive to outliers. If your data does not have outliers, then examine the Mean Average Error (MAE), which is not as sensitive to outliers.
- **RMSLE** is the Root Mean Square Logarithmic Error. It is the ratio (the log) between the actual values in your data and predicted values in the model. Use RMSLE instead of RMSE if an under-prediction is worse than an over-prediction – where underestimating is more problematic than overestimating. For example, is it worse to forecast too much sales revenue or too little? Use RMSLE when your data has large numbers to predict and you don't want to penalize large differences between the actual and predicted values (because both of the values are large numbers).