Intuitive explanations to Data Science questions

(From stackexchange and crossvalidated)

Purpose of the Link function in GLM

https://stats.stackexchange.com/questions/48594/purpose-of-the-link-function-in-generalized-linear-model

Difference between Logit and Probit model

https://stats.stackexchange.com/questions/20523/difference-between-logit-and-probit-models?noredirect=1&lq=1

What's the difference between terms 'link function' and 'canonical link function' for GLM?

https://stats.stackexchange.com/questions/40876/what-is-the-difference-between-a-link-function-and-a-canonical-link-function?rq=1

Down sampling vs Up sampling on the significance of predictors in Logistic regression

https://stats.stackexchange.com/questions/199230/downsampling-vs-upsampling-on-the-significance-of-the-predictors-in-logistic-reg

Making sense of PCA, Eigen vectors and Eigen values

https://stats.stackexchange.com/questions/2691/making-sense-of-principal-component-analysis-eigenvectors-eigenvalues

Intuitive way to understand PCA

https://math.stackexchange.com/questions/1146/intuitive-way-to-understand-principal-component-analysis

What is the intuition behind SVD?

https://stats.stackexchange.com/questions/177102/what-is-the-intuition-behind-svd/179042

Relationship between SVD and PCA?

https://stats.stackexchange.com/questions/134282/relationship-between-svd-and-pca-how-to-use-svd-to-perform-pca?noredirect=1&lq=1

What are the differences between Factor Analysis and PCA?

https://stats.stackexchange.com/questions/1576/what-are-the-differences-between-factor-analysis-and-principal-component-analysi

How to reverse PCA and reconstruct original variables from several principal components?

https://stats.stackexchange.com/questions/229092/how-to-reverse-pca-and-reconstruct-original-variables-from-several-principal-com?noredirect=1&lq=1

PCA and proportion of variance explained

https://stats.stackexchange.com/questions/22569/pca-and-proportion-of-variance-explained?noredirect=1&lq=1

Why is accuracy not the best measure of assessing classification models?

https://stats.stackexchange.com/questions/312780/why-is-accuracy-not-the-best-measure-for-assessing-classification-models

Is accuracy an improper scoring rule in a binary classification setting?

https://stats.stackexchange.com/questions/359909/is-accuracy-an-improper-scoring-rule-in-a-binary-classification-setting

How to calculate AUC or the c-statistics by hand

https://stats.stackexchange.com/questions/145566/how-to-calculate-area-under-the-curve-auc-or-the-c-statistic-by-hand?noredirect=1&lq=1

Bayesian and Frequentist reasoning in English

https://stats.stackexchange.com/questions/22/bayesian-and-frequentist-reasoning-in-plain-english

Way to remember the definitions of Type 1 and Type 2 errors

https://stats.stackexchange.com/questions/1610/is-there-a-way-to-remember-the-definitions-of-type-i-and-type-ii-errors?noredirect=1&lq=1

Why is ANOVA equivalent to Linear Regression?

https://stats.stackexchange.com/questions/175246/why-is-anova-equivalent-to-linear-regression

What is the relationship between ANOVA to compare means of several groups and ANOVA to compare nested models?

https://stats.stackexchange.com/questions/315979/what-is-the-relationship-between-anova-to-compare-means-of-several-groups-and-an

If the t-test and the ANOVA for two groups are equivalent, why aren't their assumptions equivalent?

https://stats.stackexchange.com/questions/1637/if-the-t-test-and-the-anova-for-two-groups-are-equivalent-why-arent-their-assu?rq=1

How does a Support Vector Machine (SVM) work?

https://stats.stackexchange.com/questions/23391/how-does-a-support-vector-machine-svm-work

Why should we use t errors instead of normal errors?

https://stats.stackexchange.com/questions/120776/why-should-we-use-t-errors-instead-of-normal-errors?noredirect=1&lq=1

Difference between Cross Entropy and KL Divergence

https://stats.stackexchange.com/questions/357963/what-is-the-difference-cross-entropy-and-kl-divergence?noredirect=1&lq=1

Maximum likelihood function (MLE) in Layman terms

https://stats.stackexchange.com/questions/112451/maximum-likelihood-estimation-mle-in-layman-terms

Drawbacks of K means

https://stats.stackexchange.com/questions/133656/how-to-understand-the-drawbacks-of-k-means

Pearson vs Spearman vs Kendall Correlation coefficients

https://datascience.stackexchange.com/questions/64260/pearson-vs-spearman-vs-kendall/64261

Is there an Intuitive explanation why multicollinearity is a problem in linear regression?

https://stats.stackexchange.com/questions/1149/is-there-an-intuitive-explanation-why-multicollinearity-is-a-problem-in-linear-r

Regression coefficients that flip sign after including other predictors

https://stats.stackexchange.com/questions/1580/regression-coefficients-that-flip-sign-after-including-other-predictors

What is Entropy and Information gain?

https://stackoverflow.com/questions/1859554/what-is-entropy-and-information-gain

Why does a 95% Confidence Interval not imply a 95% chance of containing the mean?

https://stats.stackexchange.com/questions/26450/why-does-a-95-confidence-interval-ci-not-imply-a-95-chance-of-containing-the

Intuitive explanation for dividing by n-1 when calculating Standard deviation

https://stats.stackexchange.com/questions/3931/intuitive-explanation-for-dividing-by-n-1-when-calculating-standard-deviation

Why is Euclidean distance not a good metric in high dimensions?

https://stats.stackexchange.com/questions/99171/why-is-euclidean-distance-not-a-good-metric-in-high-dimensions?rq=1

GBM vs XGBOOST? Key difference?

https://datascience.stackexchange.com/questions/16904/gbm-vs-xgboost-key-differences

Intuitive explanation of how UMAP works compared to t-sne?

https://stats.stackexchange.com/questions/402668/intuitive-explanation-of-how-umap-works-compared-to-t-sne

What does it mean that a Statistic T(X) is sufficient for a parameter?

https://stats.stackexchange.com/questions/84226/what-does-it-mean-that-a-statistic-tx-is-sufficient-for-a-parameter?noredirect=1&lq=1

Cross validation vs Train validation test

https://stats.stackexchange.com/questions/410118/cross-validation-vs-train-validation-test

What is the difference between test set and validation set?

https://stats.stackexchange.com/questions/19048/what-is-the-difference-between-test-set-and-validation-set?noredirect=1&lq=1

Why only three partitions (training, validation, test)

https://stats.stackexchange.com/questions/9357/why-only-three-partitions-training-validation-test?noredirect=1&lq=1