A fundamental tool in statistical modeling and machine learning are generalized linear models, or GLMs. Various frameworks, suitable for various use cases ranging from large-scale distributed computing to small-scale statistical analysis, optimize GLMs using a variety of techniques. Depending on the size of the dataset, the computing environment, and the modeling needs, the framework and optimization technique selected can have a substantial impact on accuracy, scalability, and performance.

The following table summarizes the optimization methods used by six popular packages/frameworks for fitting GLMs, along with a brief explanation of when each offers superior performance:

| **Module/framework/package** | **Name and a brief description of the algorithm** | **An example of a situation where using the provided GLM implementation provides superior performance compared to that of base R or its equivalent in Python** |
| --- | --- | --- |
| **Base R (stats::glm)** | GLM parameters are estimated using iteratively reweighted least squares (IRLS) for maximum likelihood. | Fast and easy to understand, ideal for small to intermediate datasets on a single system. |
| **High-performance R** | Fits GLMs on massive datasets or in environments with limited memory by using parallel packages (bigmemory, parallel, bigglm). | while using parallelized computation or datasets larger than RAM to fit GLMs. |
| **Dask-ML** | Dask array-based parallelized coordinate descent for GLM optimization that works with the scikit-learn interface. | Effective for huge datasets dispersed among multiple core environments or clusters. |
| **SparkR** | Spark's data-parallel engine and resilient distributed datasets (RDDs) are used in distributed IRLS. | Perfect for enterprise Spark infrastructure on large data platforms. |
| **Spark Optimization (MLlib)** | supports stochastic gradient descent (SGD), LBFGS, and mini-batch gradient descent for learning that is scalable. | Beneficial for sparse or large-scale data training of logistic regression models. |
| **Scikit-learn** | LBFGS or coordinate descent with automatic hyperparameter adjustment that accommodates L1/L2 regularization. | Excellent for high-dimensional issues such as regularized regression and text categorization. |