In his paper ” Statistical Modeling: The Two Cultures. Statistical Science” Leo Breinman discusses challenges of traditional statistical methods. He argues that stochastic data model, meaning basically random, fails to capture the complexity of phenomena it tries to describe.

Even when using best of fit and residual analysis methods might produce erroneous results or not entirely explain the underlying phenomena.

“Misleading conclusions may follow from data models that pass goodness-of-fit tests and residual checks. But published applications to data often show little care in checking model fit using these methods or any other.” (Leo Breinman, ” Statistical Modeling: The Two Cultures. Statistical Science”,p 6)

While traditional statistical inference methods may still be valid in some cases to both predict and describe data they were often favored due to their interpretability. It is much easier to understand the output from logistic regression compared to machine learning algorithms, not to mention neural nets.

The second approach, an algorithmic, black box, approach, offers often better accuracy. It often manage to describe the studied phenomena’s more accurately than traditional methods.

The article was written quite some time ago and those ideas are hardly controversial any longer. Machine learning has been adopted by Statistics and algorithmic predictive methods are part of curricula at many Statistical institutes.

However that transition is far from complete. At my work place for example, we still heavily rely on GLM even though Machine learning is slowly being considered as alternative.