Анализ данных и машинное обучение в авиации

С развитием информационных технологий, объем обрабатываемых данных непреклонно растет. Объем информации стал настолько велик, что обработка всей информации стандартными програмно-аппаратными средствами стал крайне неэффективен. Другими словами, большие данные (Big Data) – стал проблемой. По большей части, анализа и обработки данных. Проблема хранения и структурирования рассматриваться в данной статье не будет.

Под термином интеллектуального анализа данных подразумевается конкретный набор методик и технологий, повышающих эффективность обработки данных. На данный момент, исходя из предложений на рынке, основным средством является система распределенных вычислений, где обработка больших объемов информации требует для себя кластер высокопроизводительных вычислительных систем. Такие системы используют следующие методы анализа:

* Методы класса Data Mining – “обучение” системы некоторым ассоциативным правилам, принципам и категориям новых данных на основе уже проанализированных данных. Сюда же относятся кластерный и регрессионный анализ;
* Краудсорсинг – использование вычислительных мощностей некоторого круга привлеченных на основании публичной оферты, но не связанных трудовым договором;
* Машинное обучение – использование обучаемых моделей (с учителем и без), основанных на статистическом анализе и извлекающих знания из конкретного набора данных;
* Искуственные нейронные сети, генетические алгоритмы;
* Имитационное моделирование;
* И др.;

В данной статье, рассматриваются такие методы анализа больших данных как “Data mining” (рус. интеллектуальный анализ данных) и “Machine learning” (рус. машинное обучение), из-за их активного развития и эффективности.