Top 10 - on DATA

Most important things in Big Data Analysis

1. **Ensure LOTS OF useful quality data**, either by having it or getting it.
2. **Fast computing and access to data** for quick analysis turn-around is key.
3. **All data is flawed**. Make sure you know how and filter it first.
4. **Check and understand input data** and its most basic relations.
5. **Devise a good "measure of goodness"** to track improvements.
6. **Test several methods**, as different methods apply to different cases.
7. **Consider number of variables**, and omit those which don't contribute.
8. **Do not (unknowingly) overtrain methods**. Ensure this by splitting data.
9. **Don't expect miracles from Big Data**. Typical improvements are 5-20%.
10. **Make cross checks**. All data analysis results can be bugged / flawed.

Top 10 - on ALGORTIHMS

Most important things in Big Data Analysis

1. **Ensure large data samples** and understanding of it.
2. **Think carefully about the loss function**, i.e. what you want to optimise.
3. **Tree algorithms are good for getting fast results** on structured data.
4. **Neural networks are more performant and versatile**, but harder to train.
5. **Variable transformation** is typically required for Neural Networks.
6. **Image analysis** is mostly done with a Convolutional Neural Network.
7. **Dimensionality reduction** benefits high dimensional problems.
8. **Streams of data** (e.g. text) can be analysed with LSTM/RNN networks.
9. **Unsupervised learning/clustering** results can be difficult to interpret.
10. **Uncertainties in regression** can be given by ML (typically NN) algorithms.