1. **Why feature scaling influence?**  
   There's a word in applying machine learning algorithm, 'garbage in, garbage out'. The more real reflection of your features, the more accuracy your algorithm will get. That applies too for how machine learning algorithms treat relationship between features. Different from human's brain, when machine learning algorithms do the classify for example, all the features are expressed and calculated by the same coordinate system, which in some sense, **establish a priori assumption** between the features(not really reflection of data itself). And also the nature of most algorithms is to find the most appropriate weight percentage between the features to fittest the data. So when these algorithms' input is unscaled features, large scale data has more influence on the weight. Actually it's not the reflection of data itself.
2. **Why usually feature scaling improve the accuracy?**  
   The common practice in unsupervised machine learning algorithms about the parameter selection is that you should not add any personal subjective assumption about data. The best way is just to assume that they have the equality probability to appear. I think it applies here too. The feature scaling just try to make the assumption that all the features has the equality opportunity to influence the weight, which more really reflects the information/knowledge you know about the data. Commonly also result in better accuracy.