**Hierarchical Classification**

* **Flat approach**: the labels are just the end nodes
* **Local approach**: In essence, in this top-down approach, for each new example in the test set, the system first predicts its first-level (most generic) class, then it uses that predicted class to narrow the choices of classes to be predicted at the second level (the only valid candidate second-level classes are the children of the class predicted at the first level), and so on, recursively, until the most specific prediction is made
  + As a result, a disadvantage of the top-down class-prediction approach (which is shared by all the three types of local classifiers discussed next) is that an error at a certain class level is going to be propagated downwards the hierarchy, unless some procedure for avoiding this problem is used

http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=7C0EC326B47A60B9159A6BF00B4AD105?doi=10.1.1.183.302&rep=rep1&type=pdf