

Lec 13

9 декабря 2024 г. 17:08

1. Decision trees can be used as classifier or as regressor. But more often as classifier
2. Output are probabilities
3. Very predictable model, cause we know how it decides
4. Entropy is basically the measure of chaos. If there is no dominant class in S subset, than it's entropy is high, cause we are uncertain of the class of random record.
5. Otherwise if there is a dominant class, the entropy is low, cause we are certain that we are gonna pull random record and it will highly likely have this class.
6. $H = - \sum (p(c) * \log_2(p(c)))$ - always non-negative
7. There is also an $H_{attr} = - \sum (p(v) * \sum (p(c | v) * \log_2(p(c | v))))$ - sum entropy for every value of attribute attr
8. $InfoGain = H - H_{attr}$ - that shows us how this attr lower the entropy. On each step we pick an attribute that lowers the entropy the most.
9. We stop when we run out of attributes or when we reach desired height of tree. Too high trees usually are overfitted and can't be reliable on new data.