# Solutions to the multiple value issue:

## One hot encoding:

**Positive:** simple solution

**Flaw:** treats a long thread of values the same way as a short one

## Values to features:

**Positive:** reflective weighting

**Flaw:** big chance of losing data in the **cleaning process. (Although filling nan with false can fix that) + increases the size of the database.**

## Values count:

**Idea:** replace the thread with a count of it’s values.

**Positive:** simple

**Flaw:** assumes all values have the same weight

## Values to features -> evaluation-> sum:

**Idea:** turn the values to features in a new table that includes the target. Calculate each new feature’s importance, replace the thread in the database with the sum of the importance of it’s values.

**Positive:**

**flaw:** complicated+ doesn’t account for communicative relationships between the values and the rest of the features + unpredictable results I don’t know of cases using this method.