Feature distance & clustering