

## Topic modeling in court rulings

**Juan Diego Cuenca Camacho**

PhD candidate, Department of Economics, Universitat Politècnica de València, Spain.

---

### ***Abstract***

*Judges usually have to deal with the valuation of a company's shares in the event of bankruptcy, merger and acquisitions and other disputes (DiGabriele, 2006). This paper aims to take a first step in answering the question of what is the treatment of company valuation in Spanish jurisprudence. We will construct a classifier of court rulings that allows us to discriminate rulings that effectively deal with company valuation from those that do not. Three unsupervised models are proposed: Latent Dirichlet Allocation (LDA), Latent Semantic Indexing (LSI) and Nonnegative Matrix Factorization (NMF). LDA identifies two topics that have a interpretation aligned with the classification that a human being would give. Topics from LSI are not easy to interpret, whereas results from NMF are closer to the LDA ones. To estimate the goodness-of-fit of the models and compare them we use the coherence measure. According to it, LDA gets the highest score. However, for all models the number of topics that maximize coherence are greater than two, what highlights that, despite the usefulness of objective measures of topic modeling evaluation, human judgment may be more appropriate for topic modeling and model comparison.*

***Keywords:*** *Company valuation; Court ruling; Topic modeling.*

---