



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Analysis and reproducibility of paper “A Word Embedding based Generalized Language Model for Information Retrieval”

D. Ganguly, D. Roy, M. Mitra, G.J.F. Jones

11 february 2020

Alessio Lazzaron, Matteo Marchiori, Andrea Oriolo, Fabio Piras



1 Aspects Reproduced



- Lucene... What version?
- Indexing... How?
- Base Language Model

- First approach: let's do it
ourselves (fail)
- Second approach: let's try to
take example and make it
better
- Is it enough? It's never
enough. Time constraints...





Final observations





Eco, Umberto (1977), *Come si fa una tesi di laurea*, Bompiani, Milano.



Knuth, Donald Ervin (1984) *The T_EXbook*, Addison-Wesley, Reading (Massachusetts).