Chapter 1

introduction

motivation , primarily problem to solve
we need jobs to match skills we need to detect skills
real motivation secondary problem that we are actually solving
how to detect skills. why it is challenging and different than usual keyword
extraction problems

Chapter 2

literature review

2.1 what are the algorithms in keyword extraction

NLP introduction The automatic analysis of
text involves a deep understanding of
natural language by machines, a reality
from which we are still very far off [**?**]. Hitherto, online information
retrieval, aggregation, and processing
have mainly been based on algorithms
relying on the textual representation of
web pages. Such algorithms are very
good at retrieving texts, splitting them
into parts, checking the spelling and
counting the number of words. When
it comes to interpreting sentences and
extracting meaningful information, however, their capabilities are known
to
be very limited. NLp in fact, requires high level symbolic capabilities(Dyer 1994) including:
1-creation and propagation of dynamic
bindings; 2-manipulation of recursive, constituent structures, acquisition
and access of lexical, semantic, and episodic memories; control of multiple
learning/process- ing modules and routing of informa- tion among such modules; grounding of basic-level language constructs (e.g., objects and actions)
in perceptual/motor experiences; representation of abstract concepts. All
such capabilities are required to shift from mere NLP to what is usually
referred to as natural language under- standing (Allen, 1987). Today, most

of the existing approaches are still based on the syntactic representation of text, a method that relies mainly on word co- occurrence frequencies. Such algorithms are limited by the fact that they can pro- cess only the information that they can see. As human text processors, we do not have such limitations as every word we see activates a cascade of semantically related concepts, relevant episodes, and sensory experiences, all of which enable the completion of complex NLP taskssuch as word-sense disam- biguation, textual entailment, and semantic role labelingin a quick and effortless way.

supervised
unsupervised
graph-based algorithms
what works and what fails

2.2   Automatic keyphrase extraction

what does it mean to tag data by people Tagging is the process of labeling web resources based on their content. Each label, or tag, corresponds to a topic in a given document. Unlike metadata assigned by authors, or by professional indexers in libraries, tags are assigned by end- users for organizing and sharing information that is of interest to them. The organic system of tags assigned by all users of a given web platform is called a folksonomy.[**?**]

Chapter 3

our approach our algorithm

our algorithm a bit of introduction
data acquisition
data cleaning
the results

Chapter 4

visualization of trends

Chapter 5

conclusion