1. General comments

These comments are made in relation to unmoderated, unapproved scores.

17 candidates sat the examination. The average mark was 62.75%. The top mark was 85%, the bottom mark was 22%. 3 candidates scored under 50%.

All candidates answered question 1 (av. 62.10%); 4 candidates answered question 2 (av. 35%); 16 answered question 3 (av. 66.55%); 11 answered question 4 (av. 75.50%); and 2 answered question 5 (av. 36.65%). In several cases, parts were not attempted, which would clearly lead to depressed scores. Also, there were several vague, general, over-brief answers that did not persuade that the candidates really understood points at issue. Marks were lost by some, due to not answering the second (typically discussion) parts of questions. One candidate indulged in what was effectively brain-dumping, writing as much as they could remember from lecture slides and leaving the examiner to pick out what might be relevant points. That is usually a bad strategy, as it does not persuade the examiner that the candidate has assimilated and integrated knowledge to be able to construct a reasonable answer, discussion or argument in relation to a question.

2. Specific comments

Question 1

b) Some candidates did not address the second part of the question, on impact.
D) Same comment on the discussion aspect.
E) Same comment on the why is it better aspect.
F) Some candidates neglected to annotate the entire sentence as was required. Several who did attempt to annotate the entire sentence apparently forgot that annotation should have also covered quotation marks, comma and full stop.
G) Some basic mistakes by some: e.g., it is hard to do POS tagging if one does not know what the tokens are.
H) Variably answered, some basic mistakes such as thinking the head of a sentence is the subject.

Question 2

a) i) and ii) In general these parts were poorly answered, with evident confusion regarding the difference between applying a TBL tagger to an unseen text, and inducing lexicon and rules from already tagged text.
B) This was also poorly answered. One candidate complained (on their script) that they had to write far too much for 1 mark for b) i). However, marks are given on the paper for a purpose, e.g., to guide candidates regarding the amount of material expected in proportion to other parts. If one thinks that too much writing is involved, this should be a warning that, e.g., a wrong approach has been adopted or other misconception is involved. This was the case in this instance, as a basic misunderstanding was demonstrated regarding how a gazetteer and a grammar are used in a rule-based system. It would not, for example, be appropriate to include “21st February, 2013” in the gazetteer, although one would be well advised to include the names of months and their variants (e.g. “Feb.”). Likewise, it does not make sense to list all full person names in a gazetteer, or worse, full names with their titles, although one might list titles in the gazetteer, as well as e.g., common first names. It does not make sense to list in the gazetteer all distances in kilometres as in “60 km”. It does make sense to list units of length such as “km”.
C) was also poorly answered, surprising given that machine learning-based NER is a core aspect of text mining.

Question 3

d) Several candidates referred to only 1 of the 2 main methods to construct a a PCFG.
F) Some answers remained rather general and vague regarding the discussion aspect.
i) ii) Some candidates did not indicate dependency arrows, therefore it was unclear to the examiners whether there was good understanding of dependency relations.
i) iii) This was generally answered well although most candidates did not achieve full marks.

Question 4

This was the highest-scoring question. One candidate achieved 20/20.

a)-c) There was an apparent lack of firm knowledge among several candidates regarding basic notions, definitions and advantages/disadvantages, with some vague and inaccurate answers.
D) Calculations were in general well tackled. Credit was given for part work and for manipulation of formulae even though an error (including e.g. failing to apply log2) may have been introduced that prevented arriving at the perfect result. Interpretation of results in some cases revealed underlying misunderstanding, for example, a calculation yielding no strong evidence of collocation was said to show such evidence as “one could not disprove the null hypothesis”.

Question 5

a)There were some over brief and vague answers so understanding could not easily be determined.