

## What do our Children Read About? Affect Analysis of Chilean School Texts

CLAUDIA MARTÍNEZ

JORGE FERNÁNDEZ

*Universidad Católica de la Santísima Concepción, Chile*

ALEJANDRA SEGURA

CHRISTIAN VIDAL-CASTRO

CLEMENTE RUBIO-MANZANO

*Universidad del Bío-Bío, UBB, Concepción, Chile*

### ABSTRACT

*We present a study of the affective character of 1<sup>st</sup> to 8<sup>th</sup> year Chilean school texts, to which we applied lexicon-based affect analysis techniques to identify 6 basic emotions (anger, sadness, fear, disgust, surprise and happiness). First, we generated a corpus of 525 documents, 18176 paragraphs and 137516 words. Then, using the affective words frequency, we built a classifier based on Emotion Word Density to detect emotions in the texts. Our results show that the predominant affective states are happiness (58%), sadness (16%) and fear (12%). The 6 basic emotions are present in most literary forms with uniform relative density except for songs, where anger is absent. Classifier performance was validated by comparing its results against the opinions of experts in the field, and its results show an above-average conformity (accuracy = 63%), above-average predictive capacity (precision = 69%) and good classifier sensitivity (recall = 80% and f-measure = 93%).*

**Keywords:** Affect analysis, sentiment analysis, subjectivity analysis

## 1. INTRODUCTION

Affect analysis is a branch of natural language processing focused on the development and application of techniques that estimate the emotive aspects of a text [1]. Several psychological theories commonly used in affect analysis classify emotions into between six and nine basic classes. For example, Ekman [2] defines the six basic emotions as anger, disgust, fear, joy, sadness and surprise. In [3], Plutchik and Kellerman add guilt, interest and shame to the basic emotion classes. These approaches assume that all other emotions can be assigned to these basic groups, on which other theories and classifications are based. This study aims to detect emotions in school texts provided by the Chilean Education Ministry for 2014 based on a lexical resource.

The rest of the article is organized as follows: in the next section, it reviews the relevant bibliography about lexicon-based affect analysis techniques. The next two sections describe the various techniques, tools and metrics used in this study. Experiments and their results are described in sections 5 and 6, while section 7 and 8 present the expert validation stage and a discussion of the results.

## 2. BACKGROUND

Affect analysis is usually applied using an approach based on machine learning, where automated or supervised learning algorithms detect and classify texts into affective categories. Another common approach uses lexical structures, specifically lexicon containing sets of words categorized according to their affective content [4], [5].

Affect analysis and sentiment analysis have been applied to literary forms to gather information about the emotions that these texts might provoke the reader. Most works that have applied the lexicon-based strategy to this purpose have taken advantage of the availability of English-language lexicons. For example, Neviarouskaya in [6] applied the ATtitude Analysis Model to the classification of fairy tales using the Izard's theory of 9 emotions