



UNIVERSIDAD DEL BÍO-BÍO, CHILE  
FACULTAD DE CIENCIAS EMPRESARIALES  
Departamento de Sistemas de Información

INCORPORACIÓN DE INTENSIDAD AFECTIVA EN  
FORMA AUTOMÁTICA EN LEXICÓN, EMPLEANDO LA  
CLASIFICACIÓN DE PLUTCHIK PARA LA MEJORA DEL  
ANÁLISIS DE AFECTOS EN TEXTOS EN INGLÉS

TESIS PRESENTADA POR CARLOS MOLINA BELTRÁN.  
PARA OBTENER EL GRADO DE MAGÍSTER EN CIENCIAS DE LA COMPUTACIÓN  
DIRIGIDA POR: CLEMENTE RUBIO Y ALEJANDRA SEGURA

# Abstract

*Nowadays, one of the approaches to perform the affective analysis of texts is based on the use of lexicons, where the results of the analysis are directly related to the quality and magnitude of this lexical resource.*

*This work proposes the enrichment of an existing lexicon of affects, selected from those found in the systematic review of the literature, through a model that allows the incorporation of a measure that indicates the intensity of the words of each affective class.*

*Currently, the lexicons used in the affect analysis only include the words that belong to each affective class, and do not contemplate a measure that indicates the affective intensity of each word. It is considered that the inclusion of information about the intensity of the words, allows a more exhaustive analysis of the texts from the affective point of view.*

*To evaluate the lexicon with the inclusion of the measure of affective intensity, it compares the quality of the affective analysis of a previous study that uses a lexicon without affective intensity, with the lexicon enriched in this work.*

# Resumen

Actualmente, uno de los enfoques para realizar el análisis afectivo de textos se basa en el uso de lexicones, donde los resultados del análisis están directamente relacionados con la calidad y completitud de este recurso léxico. Este trabajo propone el enriquecimiento de un lexicón de afectos ya existente, seleccionado desde los encontrados en la revisión sistemática de la literatura, mediante un modelo que permita la incorporación automática, de una medida que indica la intensidad de las palabras de cada clase afectiva. Actualmente, los lexicones utilizados en el análisis de afectos solo incorporan las palabras que pertenecen a cada clase afectiva, y no contemplan una medida que indique la intensidad afectiva de cada palabra. Se considera que la inclusión de información acerca de la intensidad de las palabras, permite un análisis más exhaustivo de los textos desde el punto de vista afectivo. Para evaluar el lexicón con la inclusión de la medida de intensidad afectiva, se compara la calidad del análisis afectivo de un estudio previo que utiliza un lexicón sin intensidad afectiva, con el lexicón enriquecido en este trabajo.

# Índice General

<b>CAPÍTULO 1: INTRODUCCIÓN .....</b>	<b>1-1</b>
<b>CAPÍTULO 2: OBJETIVOS E HIPÓTESIS DE INVESTIGACIÓN.....</b>	<b>2-3</b>
2.1. HIPÓTESIS .....	2-3
2.2. OBJETIVOS.....	2-3
2.2.1. <i>Objetivo General</i> .....	2-3
2.2.2. <i>Objetivos Específicos</i> .....	2-3
2.3. ALCANCES DE LA INVESTIGACIÓN .....	2-3
2.4. METODOLOGÍA DE TRABAJO .....	2-4
<b>CAPÍTULO 3: MARCO TEÓRICO Y TRABAJOS RELACIONADOS ....</b>	<b>3-5</b>
3.1. CONCEPTUALIZACIÓN .....	3-5
3.1.1. <i>Clasificaciones Afectivas</i> .....	3-5
3.1.2. <i>Análisis Afectivo</i> .....	3-6
3.1.3. <i>Wordnet</i> .....	3-6
3.1.4. <i>Métricas de Similitud</i> .....	3-7
3.2. REVISIÓN SISTEMÁTICA DE LA LITERATURA.....	3-10
3.2.1. <i>Preguntas de Investigación</i> .....	3-10
3.2.2. <i>Protocolo de Búsqueda</i> .....	3-11
3.2.3. <i>Protocolo de Revisión</i> .....	3-11
3.2.4. <i>Selección de Estudios Primarios</i> .....	3-12
3.2.5. <i>Selección de Estudios Secundarios</i> .....	3-13
3.3. LEXICONES PARA ANÁLISIS DE AFECTOS.....	3-13
3.4. LEXICONES DE AFECTOS ETIQUETADOS CON INTENSIDAD AFECTIVA .....	3-14
3.5. CONCLUSIÓN DEL CAPÍTULO.....	3-16
<b>CAPÍTULO 4: INCORPORACIÓN DE UNA MEDIDA DE INTENSIDAD AFECTIVA EN UN LEXICÓN DE AFECTOS .....</b>	<b>4-18</b>
4.1. SELECCIÓN DE LEXICÓN AFECTIVO .....	4-21
4.2. FILTRO POR PADRES AFECTIVOS .....	4-21
4.3. SELECCIÓN DE SUBCONJUNTO DE ANÁLISIS.....	4-2
4.4. ENCUESTA DE OPINIÓN SOBRE INTENSIDAD AFECTIVA .....	4-3
4.5. MODELO DE INTENSIDAD AFECTIVA .....	4-4
4.5.1. <i>Etapa de homologación</i> .....	4-5
4.5.2. <i>Etapa de cálculo</i> .....	4-5
4.5.3. <i>Etapa de ordenamiento</i> .....	4-5

4.6. CORRELACIÓN ENTRE EXPERTOS V/S MODELO.....	4-6
4.7. CORRELACIÓN ENTRE [29] V/S MODELO .....	4-8
4.7.1. <i>Corpus de Evaluación</i> .....	4-8
4.7.2. <i>Detalle de Resultados</i> .....	4-9
4.8. CONCLUSIÓN DEL CAPÍTULO .....	4-14
4.8.1. <i>Resultados generales del modelo de intensidad afectiva.</i> .....	4-15
<b>CAPÍTULO 5: EVALUACIÓN: ANÁLISIS DE AFECTOS UTILIZANDO UN LEXICÓN ENRIQUECIDO.....</b>	<b>5-18</b>
5.1. DISEÑO DEL EXPERIMENTO.....	5-18
5.2. EXPERIMENTO DE EVALUACIÓN .....	5-20
5.3. DETALLE DE RESULTADOS DE EVALUACIÓN DE LEXICÓN ENRIQUECIDO .....	5-22
5.4. CONCLUSIÓN DEL CAPÍTULO .....	5-24
<b>CAPÍTULO 6: DISCUSIÓN .....</b>	<b>6-26</b>
6.1. DISCUSIÓN DE LOS RESULTADOS DEL MODELO DE INTENSIDAD AFECTIVA AUTOMÁTICA .....	6-26
6.2. DISCUSIÓN DE LOS RESULTADOS DEL EXPERIMENTO DE EVALUACIÓN DEL LEXICÓN ENRIQUECIDO .....	6-27
<b>CAPÍTULO 7: TRABAJO FUTURO .....</b>	<b>7-29</b>
<b>CAPÍTULO 8: CONCLUSIONES.....</b>	<b>8-30</b>
8.1. VERIFICACIÓN DE HIPÓTESIS Y OBJETIVOS.....	8-30
8.2. CONCLUSIONES GENERALES .....	8-32
<b>ANEXO A. REVISIÓN SISTEMÁTICA DE LA LITERATURA .....</b>	<b>8-37</b>
<b>ANEXO B. ENCUESTA DE OPINIÓN SOBRE INTENSIDAD AFECTIVA   8-41</b>	
B.1. ENTORNO DE DESARROLLO .....	8-41
<i>Características del Hardware</i> .....	8-41
<i>Plataforma de desarrollo</i> .....	8-41
<i>Base de Datos</i> .....	8-42
B.2. FUNCIONAMIENTO DE LA APLICACIÓN ENCUESTA DE OPINIÓN .....	8-43
B.3. RESULTADOS DE LA ENCUESTA DE OPINIÓN.....	8-44
<b>ANEXO C. IMPLEMENTACIÓN .....</b>	<b>8-46</b>
C.1. APLICACIÓN PARA EL CÁLCULO DE LA INTENSIDAD AFECTIVA .....	8-46
<i>Características del hardware</i> .....	8-46

<i>Entorno de desarrollo .....</i>	<i>8-46</i>
<i>Implementación del cálculo de intensidad afectiva.....</i>	<i>8-47</i>
<i>Resultados al ejecutar la aplicación.....</i>	<i>8-47</i>
C.2. APLICACIÓN PARA EXPERIMENTO DE EVALUACIÓN .....	8-48
<i>Entorno de desarrollo .....</i>	<i>8-48</i>
<i>Funcionamiento de la aplicación experimento de evaluación.....</i>	<i>8-49</i>
<i>Implementación de la aplicación de análisis afectivo .....</i>	<i>8-51</i>
<i>Presentación de resultados .....</i>	<i>8-52</i>
<b>ANEXO D. ANEXO MATRIZ Y ANÁLISIS KAPPA. ....</b>	<b>8-53</b>

# Índice de Figuras

FIGURA 3-1: RUEDA DE EMOCIONES DE PLUTCHIK. ....	3-6
FIGURA 3-2. FRAGMENTO DE UNA RELACIÓN ES-UN EN WORDNET.....	3-7
FIGURA 4-1. PROCESO PARA ENRIQUECER EL SUBCONJUNTO DE ANÁLISIS.....	4-19
FIGURA 4-2: EJEMPLO DEL PROCESO DE SELECCIÓN DE PALABRAS DEL LEXICÓN. ....	4-20
FIGURA 4-3. ESQUEMA DE LA RAÍZ AFECTIVA EN EL CÁLCULO DE LA SIMILARIDAD .....	4-22
FIGURA 4-4: ÁRBOL DE HIPERÓNIMOS EN WORDNET PARA LAS TRES ACEPCIONES DE LA PALABRA ÉXTASIS.....	4-1
FIGURA 4-5: EJEMPLO DE SIMILARIDAD ENTRE PALABRAS CON UN MISMO POS. ....	4-2
FIGURA 4-6. ENCUESTA DE OPINIÓN. <i>SCREENSHOT</i> APLICACIÓN WEB.....	4-4
FIGURA 4-7: CATEGORÍAS PARA EL ANÁLISIS DE CONCORDANCIA.....	4-6
FIGURA 4-8. COEFICIENTE DE PEARSON CON RESPECTO A LEXICÓN PROPUESTO EN [29].....	4-10
FIGURA 4-9. COEFICIENTE DE PEARSON PARA LA CLASE AFECTIVA <i>FEAR</i> . CON RESPECTO AL LEXICÓN CREADO EN [29].....	4-11
FIGURA 4-10. COEFICIENTE DE PEARSON PARA LA CLASE AFECTIVA <i>SADNESS</i> . CON RESPECTO AL LEXICÓN CREADO EN [29].....	4-13
FIGURA 4-11. COEFICIENTE DE PEARSON PARA LA CLASE AFECTIVA JOY, CON RESPECTO AL LEXICÓN CREADO EN [29].....	4-14
FIGURA 4-12. ÍNDICE DE PEARSON POR CADA MÉTRICA DE SIMILARIDAD PARA CADA CLASE AFECTIVA.....	4-16
FIGURA 4-13. ÍNDICE DE CORRELACIÓN EN PROMEDIO PARA CADA MÉTRICA DE SIMILARIDAD EN LAS 4 CLASES AFECTIVAS REVISADAS. ....	4-17
FIGURA 5-1. APLICACIÓN PARA ANÁLISIS AFECTIVO – CONFIGURACIÓN.....	5-21
FIGURA 5-2. APLICACIÓN PARA ANÁLISIS AFECTIVO - RESULTADOS.....	5-22
FIGURA 5-3. COMPARATIVA ENTRE MÉTODO DE ANÁLISIS POR CLASE AFECTIVA .....	5-25
FIGURA 5-4. COMPARATIVA DE RENDIMIENTO DE CADA MÉTRICA Y TIPO DE ANÁLISIS POR CLASE AFECTIVA.....	5-25
FIGURA 8-1. BASE DE DATOS PARA LA ENCUESTA DE OPINIÓN. ....	8-43
FIGURA 8-3. APLICACIÓN WEB, <i>SCREENSHOT</i> EXPERIMENTO DE EVALUACIÓN.....	8-49
FIGURA 8-4 AFECTIVO. TABLA DE RESULTADOS, <i>SCREENSHOT</i> APLICACIÓN WEB DE ANÁLISIS AFECTIVO.....	8-52

# Índice de Tablas

TABLA 1. RESUMEN DE CARACTERÍSTICAS DE LEXICONES SIN ETIQUETADO DE INTENSIDAD REVISADOS .....	3-17
TABLA 2. PALABRAS RESULTANTES DEL FILTRO APLICADO A LEXICÓN EMOLEX .....	4-1
TABLA 3. TÉRMINOS SELECCIONADOS PARA LA ENCUESTA Y ETAPA DE VALIDACIÓN .....	4-3
TABLA 4. CONCORDANCIA E ÍNDICE KAPPA POR CLASE AFECTIVA .....	4-7
TABLA 5. CORRELACIÓN DE PEARSON ENTRE RESULTADOS DE LA ENCUESTA Y RESULTADOS DEL MODELO .....	4-7
TABLA 6. COEFICIENTE DE PEARSON PARA CLASE ANGER. LEXICÓN VS PROPUESTA .....	4-9
TABLA 7. COEFICIENTE DE PEARSON PARA CLASE FEAR. LEXICÓN VS PROPUESTA .....	4-11
TABLA 8. COEFICIENTE DE PEARSON PARA CLASE <i>SADNESS</i> . LEXICÓN VS PROPUESTA .....	4-12
TABLA 9. COEFICIENTE DE PEARSON PARA CLASE <i>JOY</i> . LEXICÓN VS PROPUESTA .....	4-13
TABLA 10. ÍNDICE DE PEARSON POR CADA MÉTRICA DE SIMILARIDAD PARA CADA CLASE AFECTIVA. ....	4-16
TABLA 11. ÍNDICE PEARSON DE LA ENCUESTA EN COMPARACIÓN AL CORPUS DE EVALUACIÓN. ...	4-17
TABLA 12. EXTRACTO DE CORPUS DE TWEETS ETIQUETADOS CON AFECTO .....	5-19
TABLA 13. EVALUACIÓN DE LOS CLASIFICADORES .....	5-20
TABLA 14. MÉTRICA RECALL PARA EVALUACIÓN LEXICÓN - CLASE ANGER .....	5-22
TABLA 15. MÉTRICA RECALL PARA EVALUACIÓN LEXICÓN - CLASE FEAR .....	5-23
TABLA 16. MÉTRICA RECALL PARA EVALUACIÓN LEXICÓN - CLASE <i>SADNESS</i> .....	5-23
TABLA 17. MÉTRICA RECALL PARA EVALUACIÓN LEXICÓN - CLASE <i>JOY</i> .....	5-23
TABLA 18. ESTUDIOS INCLUIDOS Y EXCLUIDOS DE LA INVESTIGACIÓN .....	8-37
TABLA 19. FORMATO DE RESPUESTA DE ENCUESTA REAL .....	8-44
TABLA 20. ENCUESTAS POR CLASE AFECTIVA Y NIVEL DE INGLÉS .....	8-45
TABLA 21. EJEMPLO DE RESULTADOS DEL CÁLCULO SE INTENSIDAD .....	8-48
TABLA 22. MATRIZ DE RESPUESTAS A LA ENCUESTA PARA LA CLASE <i>ANGER</i> .....	8-53
TABLA 23. MATRIZ DE RESPUESTAS A LA ENCUESTA PARA LA CLASE <i>ANTICIPATION</i> .....	8-53
TABLA 24. MATRIZ DE RESPUESTAS A LA ENCUESTA PARA LA CLASE <i>DISGUST</i> .....	8-54
TABLA 25. MATRIZ DE RESPUESTAS A LA ENCUESTA PARA LA CLASE <i>FEAR</i> .....	8-54
TABLA 26. MATRIZ DE RESPUESTAS A LA ENCUESTA PARA LA CLASE <i>JOY</i> .....	8-55
TABLA 27. MATRIZ DE RESPUESTAS A LA ENCUESTA PARA LA CLASE <i>SADNESS</i> .....	8-55
TABLA 28. MATRIZ DE RESPUESTAS A LA ENCUESTA PARA LA CLASE <i>TRUST</i> .....	8-56
TABLA 29. MATRIZ DE RESPUESTAS A LA ENCUESTA PARA LA CLASE <i>SURPRISE</i> .....	8-56



# Capítulo 1: Introducción

El análisis de afectos y la minería de opinión han cobrado relevancia en los últimos años, impulsado por la explosión de las redes sociales, el comercio en línea, los blogs, entre otros. Estos medios han aumentado enormemente la cantidad de información disponible, por lo que cada vez es más necesario contar con mecanismos que permitan utilizar esta información y mejorar la toma de decisiones sobre determinados productos, servicios, personas, etc. El análisis de afectos en textos nos permite, por ejemplo, entender los afectos o emociones que evocan los candidatos a las próximas elecciones, detectar el *bullying*, comportamientos depresivos, o incluso actividad sospechosa o peligrosa en internet.

El análisis de afectos es abordado desde 2 enfoques [1], aprendizaje automático (por ejemplo redes bayesianas, máquinas de soporte vectorial, y otras), y análisis basado en lexicón, que para el análisis de afectos se entiende como una lista de palabras etiquetadas con la clase afectiva a la que pertenece.

Esta investigación se enmarca en este último enfoque y propone enriquecer un lexicón construido en base a la clasificación de 8 clases afectivas de Plutchik [2] (*Fear, Surprise, Sadness, Disgust, Anger, Anticipation, Joy y Trust*), para luego agregar a este lexicón, en forma automatizada, una medida que indique la intensidad de las palabras de cada clase afectiva. La medida de intensidad afectiva que se propone agregar tiene directa relación con el concepto *arousal* definido en [3], que se refiere a la calma o agitación que provocan los afectos. Por ejemplo, en la clasificación de afectos de Plutchik [2] para la clase afectiva *Joy* (alegría) existen los términos serenidad y éxtasis, es decir, estas palabras indican alegría, pero con distinta intensidad. Por ejemplo, si se expresa la intensidad afectiva entre 0 y 1, la palabra serenidad tendría una intensidad más cercana al 0, y éxtasis más cercano al 1.

Esta propuesta es evaluada sobre un corpus extraído de un lexicón de afectos etiquetado con intensidad afectiva, construido en idioma inglés.

Este documento se estructura de la siguiente manera: En el capítulo Objetivos e Hipótesis se formulan objetivos generales, específicos, metodología de investigación e hipótesis. El siguiente capítulo corresponde al marco teórico y trabajos relacionados, el capítulo comienza con la definición del protocolo de la

revisión sistemática de la literatura, aplicada en este trabajo. Luego se exponen los principales conceptos relacionados a la clasificación afectiva, las métricas de similaridad y los lexicones de afectos existentes, para luego exponer los trabajos que se relacionan a la construcción y/o enriquecimiento de lexicones de afectos, en específico a la estrategia/proceso utilizado para añadir una medida de intensidad afectiva. Una vez definido el marco teórico, en el capítulo siguiente, se describe el proceso para construir el lexicon enriquecido con intensidad afectiva automática y la construcción de una encuesta que permita validar el modelo que permite agregar la medida de intensidad. Luego, se describe el análisis realizado con respecto a un lexicon de afectos etiquetado con intensidad afectiva ya existente. Finalmente se describe el experimento de evaluación que permite medir el rendimiento del lexicon enriquecido automáticamente. Para terminar con los trabajos futuros y la conclusión del estudio.

## Capítulo 2: Objetivos e Hipótesis de Investigación

### 2.1. Hipótesis

Un lexicón etiquetado automáticamente con la intensidad afectiva, utilizando la clasificación de emociones de Plutchik, permite mejorar el análisis de afectos basado en lexicón para textos en inglés.

### 2.2. Objetivos

#### 2.2.1. Objetivo General

Enriquecer un lexicón afectivo basado en la clasificación de Plutchik a través del etiquetado automático de la intensidad afectiva de cada palabra utilizando *Wordnet similarity*.

#### 2.2.2. Objetivos Específicos

- Realizar una revisión sistemática de la literatura para seleccionar trabajos anteriores relevantes en el área de análisis de afectos.
- Analizar un lexicón de afectos que utilice la clasificación de emociones de Plutchik.
- Proponer un modelo que permita incorporar la intensidad afectiva en forma automática, utilizando una métrica de similaridad implementada en *Wordnet Similarity*.
- Validar la efectividad de la propuesta a partir de un experimento.
- Analizar objetivamente los resultados obtenidos y generar conclusiones y propuestas de trabajo futuro.

### 2.3. Alcances de la investigación

Aún cuando el análisis de afectos puede ser abordado desde varias perspectivas [1], este trabajo considera el análisis de afectos a través del enfoque basado en lexicón y no aborda otros enfoques como aprendizaje automático o

análisis estadístico. En conclusión, la revisión sistemática de la literatura aborda propuestas de este enfoque particular.

## 2.4. Metodología de Trabajo

Como parte de las actividades alineadas a los objetivos específicos de esta investigación se tienen:

- Realización de una revisión sistemática de la literatura, donde se recopilen los trabajos actuales de análisis de afectos basados en lexicón, junto a sus resultados y aplicaciones.
- Análisis de un lexicón de afectos que utilice la clasificación de emociones de Plutchik.
- Propuesta de modelo que permita incorporar la intensidad afectiva en forma automática, utilizando una métrica de similaridad implementada en *Wordnet Similarity*.
- Diseño de un experimento que permita validar la medida de intensidad incorporada al lexicón.
- Creación de un corpus de evaluación de análisis de afectos.
- Ejecución del experimento.
- Análisis de los resultados obtenidos.
- Preparación de conclusiones obtenidas y propuesta de trabajos futuros.

## Capítulo 3: Marco Teórico y Trabajos Relacionados

Este capítulo describe el marco teórico que guía esta investigación, el cual es recopilado mediante una revisión sistemática de la literatura. Se describe la conceptualización de elementos esenciales en el análisis de afectos basado en lexicón, se analizan las principales métricas de similaridad implementadas en *Wordnet Similarity* y, por último, el proceso de revisión sistemática de la literatura para la identificación de los trabajos relacionados.

### 3.1. Conceptualización

#### 3.1.1. Clasificaciones Afectivas

Existen 2 clasificaciones principales para definir las categorías afectivas, la clasificación de Ekman [4], y la clasificación de Plutchik [2]. Ekman propone 6 categorías: *Anger* (ira), *Disgust* (aversión, asco), *Fear* (miedo), *Happy* (alegría), *Sad* (tristeza) y *Surprise* (sorpresa). Por su parte Plutchik propone 8 categorías: *Fear* (miedo), *Surprise* (sorpresa), *Sadness* (tristeza), *Disgust* (aversión), *Anger* (ira), *Anticipation* (anticipación), *Joy* (alegría) y *Trust* (confianza), e indica que todo el resto de emociones son parte de alguna de estas clasificaciones. Plutchik organiza su clasificación en una rueda de emociones, la que se presenta en la Figura 3-1. En esta se observa que las emociones más parecidas están ubicadas más cerca, mientras las emociones opuestas están ubicadas en forma diametralmente opuestas entre ellas. El radio de la rueda indica intensidad, entre más cerca del centro mayor es la intensidad, y los espacios en blanco indican emociones complejas, formadas por las emociones básicas adyacentes.

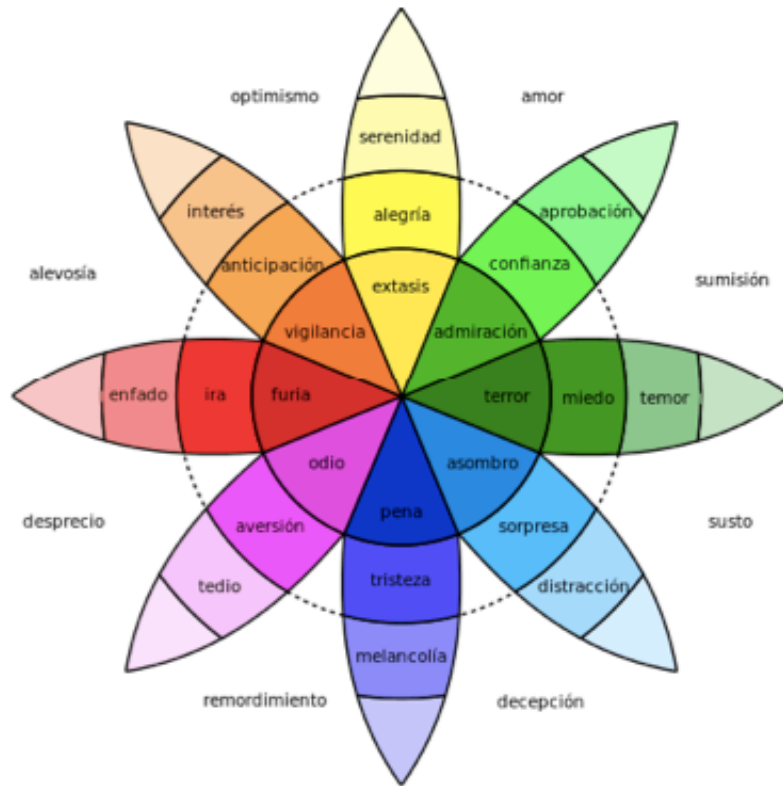


Figura 3-1: Rueda de emociones de Plutchik.

### 3.1.2. Análisis Afectivo

Dentro de las técnicas de procesamiento de lenguaje natural (PLN) basado en lexicón existen 2 que son relevantes para este estudio, el análisis de sentimientos (AS) y el análisis de afectos (AA). El AS incluye un conjunto de técnicas para extraer emociones desde textos y busca identificar opiniones, emociones y valoraciones, tanto positivas como negativas, utilizando herramientas computacionales para formalizar y polarizar este contenido [5]. Por otro lado el AA busca estimar el aspecto afectivo de un texto [6]. Un lexicón de AA contiene un conjunto de palabras clasificadas según las emociones que representa [7][8].

### 3.1.3. Wordnet

*Wordnet* [9] es una base de datos léxica ampliamente utilizada para el análisis automático de textos, que ha sido mejorada y expandida en varias investigaciones. En *Wordnet*, los sustantivos, verbos, adverbios y adjetivos están organizados por una variedad de relaciones semánticas en conjuntos de sinónimos

(*synsets*), que representan un concepto. Ejemplos de relaciones semánticas utilizadas por *Wordnet* son sinonimia, autonomía, hiponimia, miembro, similar y dominio y causa. Algunas de estas relaciones se utilizan para establecer una relación semántica, estas relaciones se asocian formando una estructura jerárquica, lo que la convierte en una herramienta útil para la lingüística computacional y el procesamiento del lenguaje natural. En la Figura 3-2 se puede ver un ejemplo de una relación en *Wordnet*. Dentro de estas mejoras se incluye un conjunto de medidas que permiten calcular la similaridad entre 2 palabras basándose en criterios como la distancia o el contenido de información. Por su parte, las medidas de similaridad que están implementadas en *Wordnet* se agrupan en el concepto *Wordnet Similarity*.

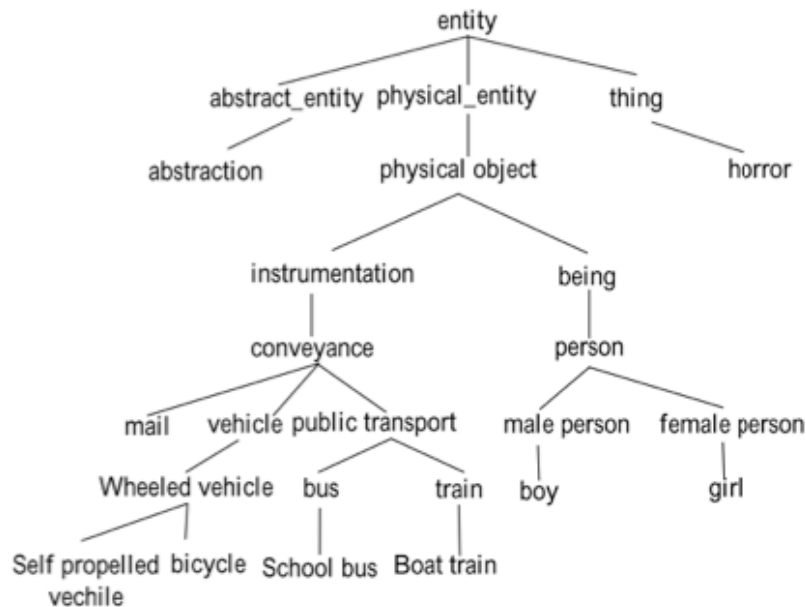


Figura 3-2. Fragmento de una relación es-un en Wordnet

### 3.1.4. Métricas de Similaridad

Para la calcular la similaridad entre dos conceptos existen 2 enfoques, el sintáctico y el semántico. El enfoque sintáctico analiza las palabras desde el punto de vista de su construcción y el enfoque semántico analiza el significado de estas, lo cual lo hace relevante para el análisis de afectos.

Dentro de las métricas de similaridad semántica que existen, hay un conjunto que están implementadas en Wordnet, a este conjunto se le llama *Wordnet Similarity*. Dentro de las medidas de similaridad, existen algunas

basadas en la estructura y otras basadas en el contenido de información (IC, por las siglas en inglés de *Information Content*) (ecuación 1). IC nos dice que mientras más información comparten dos conceptos, más similares son. Antes de explicar cada una de las medidas de similaridad, es necesario definir algunos conceptos que forman parte de cada una de ellas.

1.  $\text{lon}(c_i, c_j)$ : la longitud de la ruta más corta entre el *synset*  $c_i$  y el *synset*  $c_j$  en el árbol de *Wordnet*.
2.  $\text{lcs}(c_i, c_j)$ : el ancestro común más cercano (*lowest common subsumer* en inglés) de  $c_i$  and  $c_j$ .
3.  $\text{profundidad}(c_i)$ : la longitud de la ruta desde el *synset*  $c_i$  hasta la raíz de *Wordnet*.
4.  $P$ : la máxima  $\text{profundidad}(c_i)$  de la taxonomía.
5.  $n\_max$ : el máximo de conceptos que existen en la taxonomía.
6.  $\text{SIM}(c_i, c_j)$ : La similaridad semántica entre los conceptos  $c_i$  y  $c_j$ .

Para dos conceptos  $c_i$  y  $c_j$  comparados en una taxonomía como la representada en la Figura 3-2, la longitud del camino más corto desde el concepto  $c_i$  hasta el concepto  $c_j$  se puede determinar a partir de uno de tres casos.

- Caso 1:  $c_i$  y  $c_j$  son el mismo concepto, esto significa que  $c_i, c_j$  y  $\text{lcs}(c_i, c_j)$  es el mismo nodo. En este caso se dice que la longitud entre  $c_i$  y  $c_j$  es 0 ( $\text{lon}(c_i, c_j)=0$ ).
- Caso 2:  $c_i$  and  $c_j$  no son el mismo nodo, pero  $c_i$  es padre de  $c_j$ . Entonces el  $\text{lcs}(c_i, c_j)$  es  $c_i$ . En este caso decimos que la longitud entre  $c_i$  y  $c_j$  es 1, ( $\text{lon}(c_i, c_j) = 1$ ).
- Caso 3:  $c_i$  y  $c_j$  no son el mismo concepto, uno no es padre de otro, en este caso contamos la longitud real de la altura entre  $c_i$  y  $c_j$ , por lo tanto  $1 < \text{lon}(c_i, c_j) \leq 2 * P$

En [10] se encuentra una definición y explicación a seis de estas medidas. Entre las medidas basadas en estructura se encuentra PATH [11], que sólo se basa en  $\text{lon}(c_i, c_j)$ , y asume que  $\text{SIM}(c_i, c_j)$  se basa en la distancia entre los conceptos. La medida propuesta en [12], llamada WUP (ecuación 2) en relación a los autores, y considera la ubicación de los conceptos en la taxonomía, junto a la profundidad del LCS (*least common subsumer*).



$$IC(c_i) = -\log P(c_i)$$

Ecuación 1. Information Content (IC)

$$SIM_{WUP}(c_i, c_j) = \frac{2 * profundidad(lcs(c_i, c_j))}{profundidad(c_i) + profundidad(c_j)}$$

Ecuación 2. Métrica de similaridad de Wu&Palmer

Otra medida basada en estructura llamada LCH se encuentra en [13] (ecuación 3), donde se calcula la ruta más corta entre los dos términos, y la profundidad de la taxonomía.

$$SIM_{LCH}(c_i, c_j) = \frac{lon(c_i, c_j)}{2 * P}$$

Ecuación 3. Métrica de similaridad de Leacock & Chodoroy

Las medidas basadas en IC requieren que cada concepto incluya información adicional, ya que entre más información compartan los conceptos, más similares son. En esta categoría se encuentra la propuesta de Resnik (RES)[14] (ecuación 4), que propone que la similaridad entre 2 conceptos depende del IC de LCS entre  $c_i$  y  $c_j$ . La propuesta de Jiang & Conrath (JCN) [15] (ecuación 5), que calcula la distancia semántica para obtener la similaridad, las que son opuestas, y por último, la propuesta de Li (LIN) [16] (ecuación 6), que utiliza tanto la cantidad de información necesaria para establecer la similitud entre los dos conceptos como la información necesaria para describir completamente estos términos.

$$SIM_{RES}(c_i, c_j) = IC(lcs(c_i, c_j))$$

Ecuación 4. Métrica de similaridad de Resnik

$$SIM_{JCN}(c_i, c_j) = \frac{1}{IC(c_i) + IC(c_j) - 2 * IC(lcs(c_i, c_j))}$$

Ecuación 5. Métrica de similaridad de Jiang & Conrath

$$SIM_{LIN}(c_i, c_j) = \frac{2 * IC(lcs(c_i, c_j))}{IC(c_i) + IC(c_j)}$$

Ecuación 6. Métrica de similaridad de Lin

Es necesario señalar que existen otras medidas de similaridad en Wordnet, basadas en características y otras híbridas que utilizan la estructura y las características en su funcionamiento. Entre estas medidas está la de Tversky [17] y la de Zhou [18], estas propuestas no se consideran en este trabajo ya que incorporan descripciones de las palabras o necesitan que ciertos parámetros sean mejorados manualmente. En [19] se describe el funcionamiento de cada una de estas propuestas.

## 3.2. Revisión Sistemática de la literatura

Se realizó una revisión sistemática de la literatura con la finalidad de recopilar el estado del arte de la temática de esta investigación. Dicha revisión consistió en el análisis de 34 textos, los cuales, se revisaron e incorporaron de acuerdo a filtros y criterios de selección. De acuerdo a lo expuesto en secciones anteriores, este trabajo considera el análisis de afectos a través del enfoque basado en lexicón y no aborda otros enfoques como aprendizaje automático o análisis estadístico. A continuación, se detalla el protocolo seguido en la revisión sistemática y posteriormente sus resultados.

### 3.2.1. Preguntas de Investigación

Se formularon las siguientes preguntas de investigación que guiaron la revisión sistemática, orientadas a el análisis de afectos basado en lexicón, a buscar dentro de la literatura:

- (a) ¿El análisis de afectos se basa en el uso de lexicón o lexicones?
- (b) ¿Cuál es la clasificación afectiva del lexicón o lexicones utilizados?
- (c) ¿Cómo se construyó el lexicón utilizado en el análisis?
- (d) ¿Se utilizó *Wordnet* como fuente de construcción del lexicón?
- (e) ¿El lexicón contiene palabras existentes en *Wordnet-Affect*?

### 3.2.2. Protocolo de Búsqueda

Para realizar la búsqueda de literatura, se utilizan las siguientes cadenas de búsqueda:

(“*Document Title*”: *lexicon* AND (“*Document Title*”: *affect* OR “*Document Title*”: *affective* OR “*Document Title*”: *emotion*)) AND (“*Abstract*”: “*affect analysis*” OR “*Abstract*”: “*affective analysis*” OR “*Abstract*”: “*emotion analysis*”).

La estrategia de búsqueda se basa en los siguientes aspectos:

- En recursos con herramientas de búsqueda: Ingresar de forma escalada los términos y las combinaciones entre ellos.
- En Internet: Utilizar *Researchgate*, *Web of Science*, *Springer*, *IEEE*, *ACM* y *ScienceDirect* como motores de búsqueda, debido a que estos son los principales buscadores utilizados. Si uno o más documentos están inaccesibles, buscar en páginas y sitios alternativos.
- Autores: Identificar autores relevantes, acceder directamente a sus páginas personales para la búsqueda de material.
- En artículos: Detectar referencias bibliográficas de utilidad, en base a ellas buscar directamente el artículo citado usando los antecedentes que aparecen en la citación.

### 3.2.3. Protocolo de Revisión

El protocolo de revisión de literatura, consta de los siguientes aspectos:

(a) Normas de Revisión: Recopilación de trabajos completos e impresos. Revisión crítica de introducción, resumen, conclusión y referencias. Se decide su inclusión/exclusión respecto a normas de este protocolo. Se rotula cada artículo comentando la decisión y comentarios relevantes.

(b) Criterios de Inclusión: Se incluyen todos aquellos trabajos o estudios que abordan el tema de análisis de afectos basado en lexicón que están disponibles en portales y/o en la web que se enmarcan dentro de los siguientes temas:

- Análisis de afectos utilizando lexicones para textos en inglés.
- Construcción y/o expansión de lexicones de afectos en inglés.
- Análisis de calidad de lexicones de afectos existentes.

- Desafíos y trabajos futuros relacionados a análisis de afectos basado en lexicón.

(c) Criterios de Exclusión: Se excluyen todos aquellos estudios que, a pesar de contener los términos de búsqueda o una combinación de ellos, no contienen información relevante sobre el tema y/o no abordan el tema de interés. En esta etapa se descartan todos los estudios que no realizan análisis de afectos basado en lexicón ni describen de forma clara el experimento de evaluación.

(d) Estrategia de Extracción de Datos: Por cada estudio seleccionado, se realiza una lectura crítica con el objeto de extraer datos para el trabajo. Primero se lee la introducción, resumen, conclusión y referencias para saber:

- Introducción y Referencias: A qué comunidad está dirigido el estudio.
- Resumen, Introducción y Conclusión: Cuales son sus contribuciones y aportes.
- Resumen, Conclusión e Introducción: Cuales son las consecuencias de sus contribuciones y aportes y como se ven reflejados en aplicaciones de la vida real.
- Cuerpo del Artículo: Incluir, separar y analizar detalladamente la información útil en el estudio.
- Cuerpo del Artículo: Comprender los experimentos, el marco de trabajo sobre el cual fueron desarrollados.
- Conclusión: Observar y analizar trabajos futuros.

(e) Estrategia de Síntesis de Datos: Los datos son resumidos de acuerdo a los siguientes temas:

- Estado del arte y avances en el análisis de afectos basado en lexicón.
- Lexicones para el análisis de afectos.
- Proceso de análisis de afectos y características del experimento.
- Tipo de clasificación afectiva utilizada (Plutchik, Ekman, otra).
- Resultados del experimento.

### **3.2.4. Selección de Estudios Primarios**

Se realiza la búsqueda en los portales Web of *Science*, *Springer*, IEEE, ACM y *ScienceDirect*, además de analizar las referencias de los artículos más

relevantes. En esta etapa se recopilan 34 artículos. En el Anexo A, página 8-37, se presenta el detalle de los artículos incluidos y excluidos.

### 3.2.5. Selección de Estudios Secundarios

De los 34 estudios seleccionados en la etapa primaria, se procede a leer cada artículo en detalle, de acuerdo al protocolo de revisión descrito. Además, se desechan los trabajos que están orientados principalmente a la teoría del análisis afectivo dentro del ámbito de la psicología, debido a que este trabajo está enfocado esencialmente al estudio del análisis de afectos en términos computacionales, dejando de éstos solo los más relevantes para fines de definiciones formales. Finalmente, se seleccionan 10 trabajos relacionados a la creación de lexicones de afectos para textos en inglés, y en 2 que se relacionan a la creación de lexicones etiquetados con intensidad afectiva. Cabe destacar que en este trabajo se incluyen numerosos artículos además de los mencionados en esta revisión sistemática, ya que los conceptos tratados en esta investigación son bastante amplios y se necesitan diversas fuentes para abordarlos de forma adecuada.

## 3.3. Lexicones para análisis de afectos

Los lexicones que se utilizan normalmente en el AA contienen distintas clasificaciones de emociones básicas, y suponen que todo el resto de emociones son parte de alguna de estas clasificaciones. En [7] se propone un lexicón de afectos llamado *Wordnet-Affect*. Este lexicón se construye a partir de la base de conocimientos *Wordnet* mediante la selección y etiquetado de conceptos afectivos. *Wordnet-Affect* clasifica las palabras basándose en las seis clasificaciones afectivas de Ekman [20], cada palabra incluida en el lexicón incluye información léxica y afectiva. Luego, en [8] los autores añaden etiquetas (positiva, negativa, ambigua, neutra) que entregan información relacionada a la polaridad emocional de las palabras de *Wordnet-Affect*. A continuación, ellos proponen una variación a una métrica existente, la que permite calcular el peso afectivo de una palabra en comparación a los términos de una clase afectiva. Otro lexicón de afectos llamado *EmoLex* es presentado en [21], en este lexicón se catalogaron 14.182 palabras en 8 categorías afectivas, basándose en la clasificación afectiva de Plutchik [2]. Este lexicón se genera a partir de una lista de palabras extraídas desde *Macquarie*

*Thesaurus Wordnet-Affect* y para la clasificación se utilizó *Amazon's Mechanical Turk*, una plataforma de crowdsourcing (trabajo colaborativo). que permite realizar trabajos utilizando la inteligencia de muchos humanos, este trabajo es expandido por los autores en [22][23], donde añaden nuevos términos y agregan una estrategia de validación para asegurar la correcta clasificación de las palabras.

En [24][25] los autores construyen un lexicón para análisis de afectos de dominio específico en la red social twitter, el trabajo consiste en clasificar en 8 categorías afectivas palabras contenidas en los tweets, basados en la clasificación afectiva de Plutchik, esto detectando hashtags que se puedan asociar a emociones, algo muy similar fue presentado por otros autores en [26], sin embargo ellos no se limitan a la clasificación afectiva de Plutchik, y añaden otras clasificaciones, junto a sinónimos y polaridad. En [27] se presenta un lexicón llamado CLex, para análisis afectivo que asocia los colores a palabras y emociones, creado a partir de crowdsourcing. En total, CLex tiene 3.397 términos afectivos únicos, representando sentimientos, emociones, actitudes y comportamientos. Los términos afectivos en CLex también incluyen las 8 emociones básicas según la clasificación de Plutchik.

### **3.4. Lexicones de afectos etiquetados con intensidad afectiva**

A continuación, se presentan los estudios seleccionados en la revisión sistemática que construyen lexicones de afectos etiquetados con intensidad afectiva. Para cada uno de ellos, se muestra un resumen de características y conceptos.

En [28] los investigadores construyen un lexicón de afectos utilizando una herramienta de crowdsourcing (trabajo colaborativo) desarrollada para la investigación. Esta herramienta permite que los visitantes de un sitio web de noticias puedan evaluar la emoción principal que les entrega la noticia que acaban de leer. Con esta técnica se obtienen más de 25 mil respuestas de usuarios, lo que representa más de 13 millones de palabras disponibles para el estudio. Con la información recopilada se construye una matriz documento-emoción, que contiene los porcentajes de votación por cada documento en las 8 categorías que se definieron (Miedo, Diversión, Enojo, Irritación, No me importa, Felicidad, Inspiración, Tristeza). Una vez recopilada la información de base, y

construida la matriz documento-emoción, los investigadores proceden a etiquetar las palabras contenidas en los documentos según su tipo (sustantivo, verbo, adjetivo, etc.) y crean un filtro para mantener en el estudio sólo a las palabras presentes en *Wordnet* [9]. Con las palabras restantes se construye una matriz palabra-emoción que indica la frecuencia de la palabra en cada una de las 8 categorías afectivas. El objetivo de esta última matriz, es indicar el peso afectivo de una palabra en cada categoría afectiva. Esta matriz representa al lexicón construido, llamado *DepechMood*<sup>1</sup>, el que contiene 37 mil palabras etiquetadas con la categorización afectiva.

En [29] el investigador construye un lexicón de afectos etiquetado con una medida de intensidad afectiva, la cual indica que tan intenso es el afecto evocado por una palabra en una categoría afectiva, utilizando una escala entre 0 y 1, siendo 0 la mínima intensidad, y 1 la máxima intensidad. Se destaca que el autor utiliza un subconjunto de 4 categorías afectivas para el trabajo (*Anger, Fear, Joy, Sadness*) de las 8 definidas por Plutchik. El autor describe su proceso separándolo en 2 etapas principales:

- (a) Selección de Términos: El autor utiliza Twitter<sup>2</sup> para recopilar los términos, debido a que en esa red social se pueden encontrar gran variedad de contenidos como emociones, comentarios sobre productos y opiniones sobre muchos tópicos. Además, los *Twitters* incluyen *emojis*, hashtags y otro tipo de textos en un lenguaje no estándar. Esto aportará a que el lexicón resultante sea de amplio espectro. El autor hace hincapié en el hecho de que no es eficiente hacer un análisis de todas las palabras en todas las categorías afectivas, y que es preferible hacer listas separadas por cada categoría afectiva. Para esto indica que un término debe cumplir al menos 1 de las 2 propiedades indicadas para pertenecer a una categoría:
  - (i) Se sabe que la palabra está asociada con la emoción (aunque se desconoce la intensidad de la emoción que transmite).

---

<sup>1</sup> <https://github.com/marcoguerini/DepecheMood/releases/>

<sup>2</sup> Twitter, red social de microblogging, <https://twitter.com/>

- (ii) La palabra tiene una tendencia a aparecer en tweets que expresan la emoción.

Con las propiedades descritas anteriormente, se incluyen los términos desde 2 fuentes:

- (i) Las palabras existentes en el lexicón descrito en [23] que estén asociadas con cualquiera de las emociones de la categorización afectiva Plutchik.
  - (ii) Las palabras que tienden a coexistir más a menudo con los hashtags del lexicón palabra-emoción descrito en [24]
- (b) Anotación de intensidad afectiva de las palabras: El autor construye una encuesta que presenta 4 palabras, donde se debe elegir la que representa la mayor intensidad dentro de una categoría afectiva, y la que representa la menor intensidad. Las palabras son extraídas aleatoriamente del listado de palabras recopilado en el paso anterior, controlando que las mismas 4 palabras aparezcan un mínimo de 50 veces para ser respondidas. Para realizar la encuesta, el autor utiliza la plataforma de crowdsourcing *Crowdflower*<sup>3</sup>. El autor añade una estrategia para evitar respuestas maliciosas que puedan afectar a los resultados de la investigación, esta estrategia es denominada “Preguntas de Oro”, y consiste en preguntas que el mismo autor responde y que compara con las respuestas de los encuestados. Bajo esta técnica se establece que, si un encuestado responde correctamente menos del 70% de las “Preguntas de Oro” puede ser causal, o bien no comprendió las instrucciones, o respondió de forma mal intencionada. Estas respuestas son desechadas del estudio.

### 3.5. Conclusión del Capítulo

Mediante la revisión sistemática de literatura se analizan los principales conceptos relacionados con el análisis de afectos basado en lexicón, las clasificaciones afectivas más reconocidas y las métricas de similitud implementadas en *Wordnet Similarity*. Por otra parte, se destacan los trabajos en

---

<sup>3</sup> <https://www.crowdflower.com/>



los cuales se crean lexicones de afectos utilizando distintas estrategias de construcción. Además, se estudian trabajos relacionados al enriquecimiento de lexicones. En estos casos se revisan las técnicas de construcción, su dominio de operación y la clasificación afectiva sobre la que se construyen. La Tabla 1 resume las características comunes de todos los lexicones revisados que no están etiquetados con intensidad afectiva.

Del análisis realizado a los lexicones etiquetados con intensidad afectiva se puede determinar que si bien el primer trabajo revisado [28] construye un lexicon etiquetado con intensidad afectiva, lo hace en función de la frecuencia con que una palabra se asocia a cada clase afectiva, lo que es relevante pero no provee información de la intensidad afectiva dentro de una misma clase. El estudio revisado en [29] entrega información de la intensidad afectiva para una palabra dentro de una clase afectiva. Esto es relevante para esta investigación, ya que cumple el mismo propósito, sin embargo el método de construcción del lexicon revisado en [29] basado en el etiquetado manual y “voluntario” de los usuarios de una comunidad digital, lo que difiere del método automático que se propone en esta investigación.

Tabla 1. Resumen de características de lexicones sin etiquetado de intensidad revisados

Trabajo	Clasificación afectiva	Propósito (general/específico)	Técnica de construcción	Fuente de datos
[7]	Ekman	General	Manual	Wordnet, OpenMind
[21][22][23]	Plutchik	General	Manual	Wordnet-Affect, Google n-gram
[24][25]	Plutchik	Específico (Twitter)	Automática	Twitter
[26]	Plutchik, otras	Específico (Twitter)	Automática	Twitter
[27]	Plutchik	Específico (textos con colores)	Manual	Amazon MT

Tal como se observa en la Tabla 1, la mayor cantidad de estudios utiliza la clasificación de Plutchik y en más de la mitad de ellos, se realiza un procesamiento manual. El último trabajo reportado por el investigador referente del área fue publicado el año 2017 [29], posterior al inicio de esta investigación.

## Capítulo 4: Incorporación de una medida de intensidad afectiva en un lexicón de afectos

En términos generales la eficiencia y efectividad del análisis de afectos basado en lexicón depende:

- (i) tamaño, cantidad de palabras
- (ii) la correcta clasificación de las palabras en cada clase afectiva
- (iii) etiquetado de intensidad de las palabras en cada clase afectiva.

Cuando la construcción y/o enriquecimiento del lexicón es manual, los resultados resultan ser más efectivos, pero el proceso resulta ser más costoso en cuanto al tiempo requerido del usuario. Por su parte, el procesamiento automático es menos costoso en cuanto al tiempo, aunque por lo general es menos efectivo en los resultados.

En esta tesis se propone un proceso automático para el enriquecimiento de un lexicón afectivo, en el cual se agrega la intensidad afectiva a cada palabra basado en las métricas de similaridad implementadas en *Wordnet Similarity*.

El proceso presentado en la Figura 4-1, describe las etapas ejecutadas para enriquecer las palabras de un lexicón con intensidad afectiva.

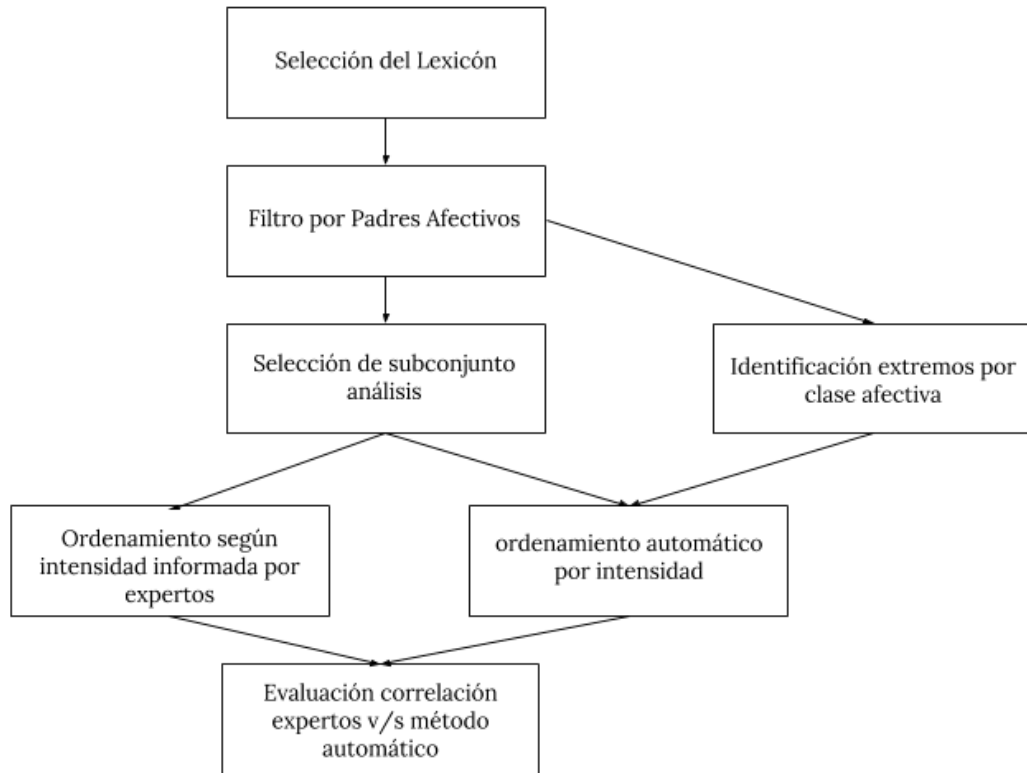


Figura 4-1. Proceso para enriquecer el subconjunto de análisis

La primera etapa consiste en seleccionar el lexicón afectivo, para esta investigación un recurso léxico en idioma inglés clasificado según Plutchik. Una vez que se cuenta con un lexicón afectivo, se deben identificar y seleccionar sólo las palabras que tengan una raíz afectiva en *Wordnet-Affect*. Posteriormente, de este subconjunto del lexicón filtrado se selecciona una muestra de estudio, y se extraen las palabras que son extremos superiores de cada clase afectiva de Plutchik, que se entiende como el término que representa la mayor intensidad afectiva en una clase, y que se encuentra más cerca del centro en la rueda de emociones de Plutchik (Por ejemplo para la clase ALEGRÍA, la palabra que evoca mayor intensidad afectiva según Plutchik es EXTASIS, ver Figura 4-2).

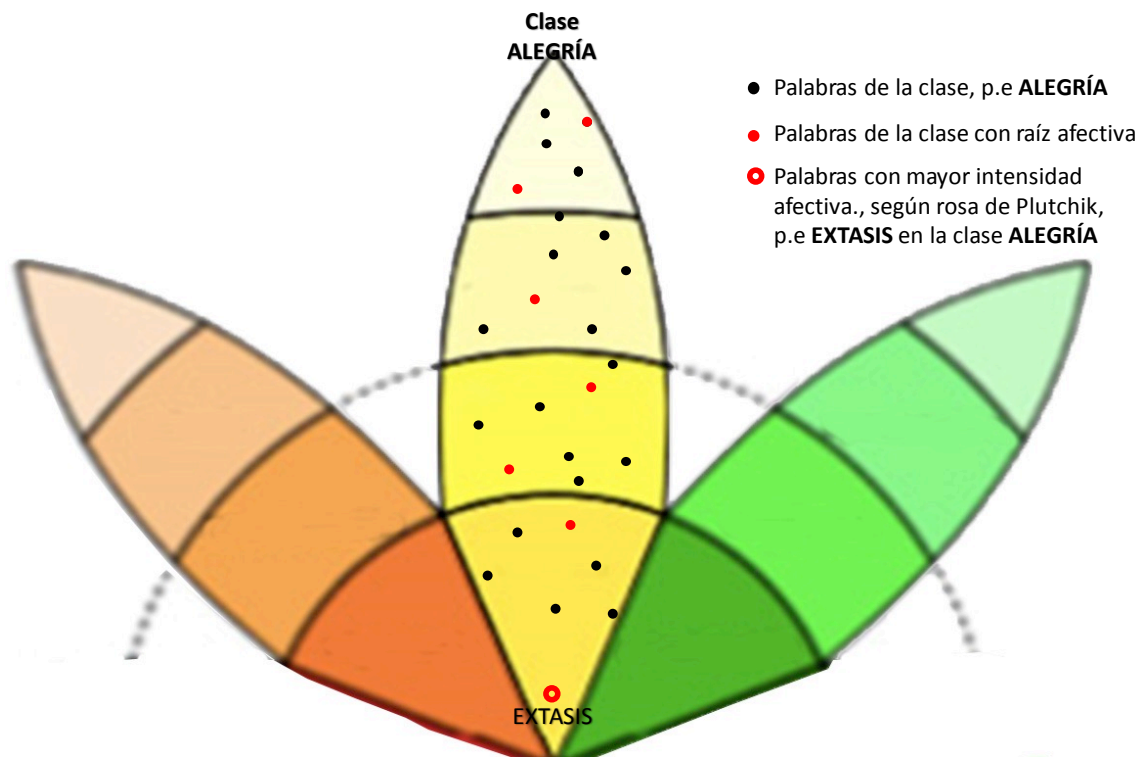


Figura 4-2: Ejemplo del proceso de selección de palabras del lexicon.

Las palabras que son extremos de una clase en Plutchik y que no tienen una raíz afectiva, se reemplazan por un término similar que cumpla la condición, esta tarea es realizada según el criterio de un experto en idioma inglés, de esta forma se desea asegurar que las palabras que reemplazan a las originales son las más adecuadas en significado y contexto.

Una vez creado el subconjunto de palabras a analizar, se consulta a un grupo de personas con estudios en idioma inglés, para que ordenen las palabras de cada clase según la intensidad de afectos que evocan.

Por otra parte, se realiza el proceso de cálculo de la intensidad afectiva de cada palabra del subconjunto según el modelo de intensidad propuesto, el cual es descrito en detalle más adelante.

Por último, se analiza la correlación entre el ordenamiento realizado por los expertos y el ordenamiento resultante de la aplicación del modelo de intensidad propuesto.

A continuación, se detalla cada una de las etapas del proceso.

## 4.1. Selección de lexicon afectivo

El lexicon afectivo que fue seleccionado para esta investigación, se ajusta a los siguientes criterios.

1. Se basa en la clasificación afectiva de Plutchik.
2. Se encuentra disponible en idioma inglés.
3. Es de propósito general.

Es importante destacar que si bien existen otras clasificaciones afectivas como la de Ekman [20], que es utilizada en otros trabajos, para esta investigación se considera que la clasificación de ocho clases afectivas propuesta por Plutchik [2] tiene un mejor balance de las emociones ya que se pueden formar 4 pares de emociones opuestas (*Joy-Sadness*, *Anger-Fear*, *Trust-Disgust*, *Anticipation-Surprise*). Además de lo anterior, Plutchik indica que las palabras evocan una intensidad dentro una clase afectiva, esto se puede visualizar en la rueda de emociones presentada en la Figura 3-1.

De acuerdo con lo anterior, el lexicon que cumple estos criterios es el propuesto en [23], llamado EmoLex. El cual es particularmente útil ya que utiliza *Wordnet-Affect* como fuente de información, lo que hace probable que tengan una raíz afectiva, y en consecuencia sea posible utilizar las métricas de similaridad.

## 4.2. Filtro por padres afectivos

Wordnet organiza las palabras como árboles según el sentido que estas tienen, es decir, una palabra podría estar en más de un árbol. Por lo tanto, para garantizar la correcta interpretación de las métricas de similaridad, se debe restringir que las palabras utilizadas pertenecen a un mismo árbol. Esto aplica tanto para las métricas basadas en estructura, como para las basadas en Contenido de Información (IC). La Figura 4-3 grafica el efecto de calcular la similaridad entre las palabras A y B cuando ambas pertenecen a un sub-árbol con la misma raíz afectiva.

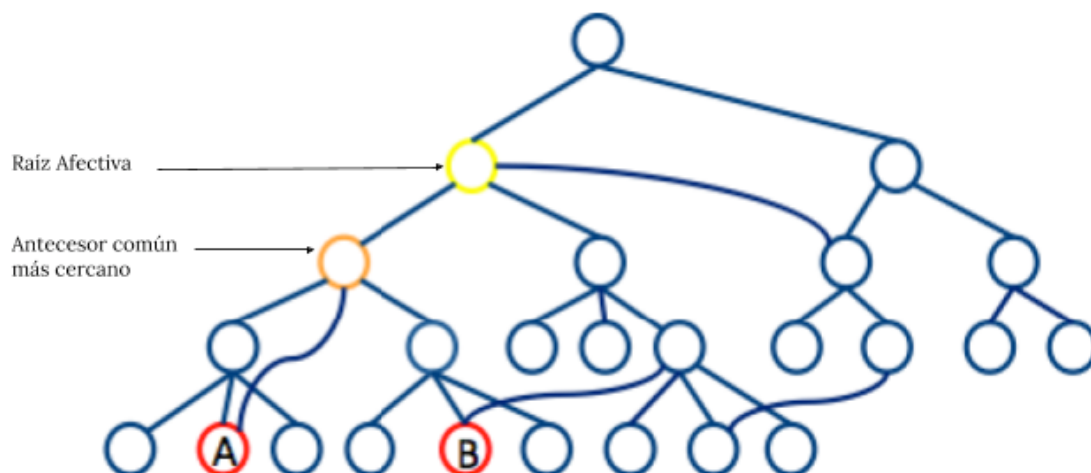
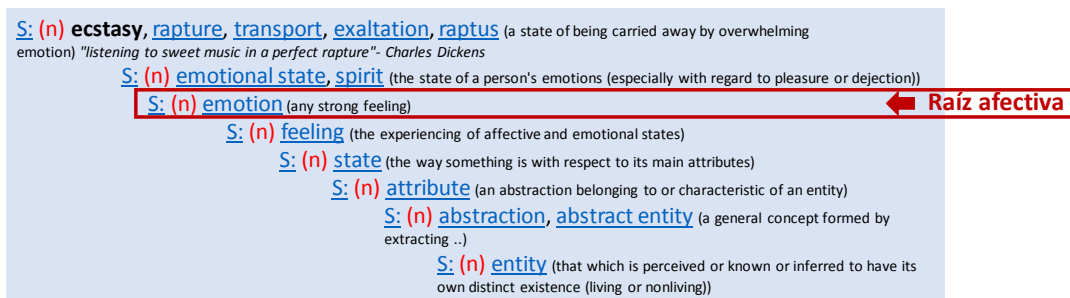


Figura 4-3. Esquema de la raíz afectiva en el cálculo de la similitud

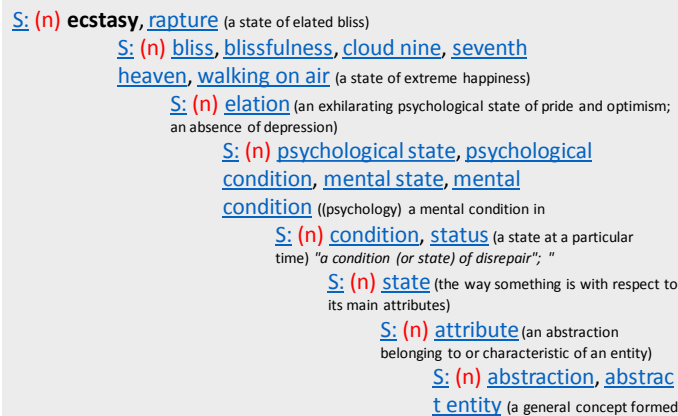
Por ejemplo, la palabra éxtasis (*ecstasy*) tiene 3 acepciones como sustantivo y para cada una de ellas tiene un árbol de padres distinto. Por lo tanto, si las palabras-acepción entre las que se compara la similitud no poseen antecesores comunes dentro del árbol de *Wordnet*, entonces es posible que los resultados de similitud generen interpretaciones erróneas.

De acuerdo con lo anterior, en este estudio, se restringen las palabras a aquellas que tienen como antecesor al menos una de las siguientes palabras {*emotion, feeling, affective, mood*}.

### Acepción 1: “emoción”



### Acepción 2 “estado mental”



### Acepción 3 “droga”

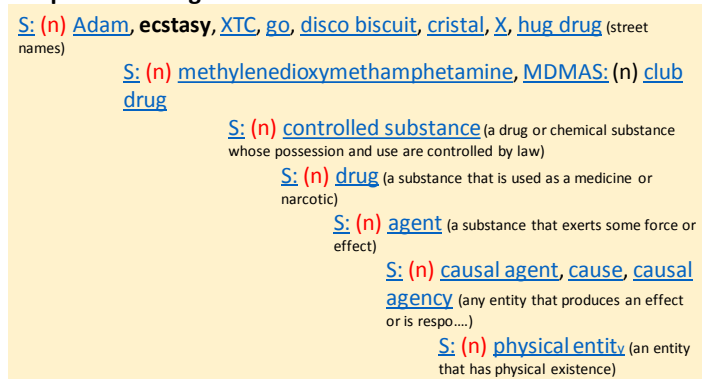


Figura 4-4: Árbol de hiperónimos en Wordnet para las tres acepciones de la palabra éxtasis

El lexicón EmoLex contiene 8.265 palabras, 1.247 para la clase *Anger*, 839 para la clase *Anticipation*, 1058 para la clase *Disgust*, 1476 para la clase *Fear*, 689 para la clase *Joy*, 1.191 para la clase *Sadness*, 534 para la clase *Surprise* y 1.231 para la clase *Trust*.

Una vez realizado el proceso de filtrado, la cantidad de términos se reduce considerablemente, llegando a 640 palabras en total, esto representa un 7,7% del lexicón EmoLex original, esta información se detalla en la Tabla 2.

Tabla 2. Palabras resultantes del filtro aplicado a lexicón EMOLEX

	EmoLex	EmoLex + Filtro	Porcentaje Restante
<i>Anger</i>	1.247	111	8,9%
<i>Anticipation</i>	839	67	8%
<i>Disgust</i>	1.058	69	6,5%
<i>Fear</i>	1.476	116	7,9%
<i>Joy</i>	689	76	11%

<i>Sadness</i>	1.191	100	8,4%
<i>Surprise</i>	534	51	9,6%
<i>Trust</i>	1.231	50	4,1%
Total	8.265	640	7,7%

Cabe mencionar que algunas de las métricas de similaridad, para ser calculadas, requieren que las palabras a comparar sean del mismo POS (*Part of Speech*), es decir, que al calcular la similaridad entre las palabras 1 y 2, ambas tengan un rol similar, ya sea como sustantivo, verbo, adverbio o adjetivo. Por ejemplo, tal como se representa en la Figura 4-5 , para calcular la similaridad entre *joy* y *ecstasy* (alegría y éxtasis) ambas palabras deben tener un antecesor afectivo con similar POS, en este caso sustantivo (*noun*, en inglés).

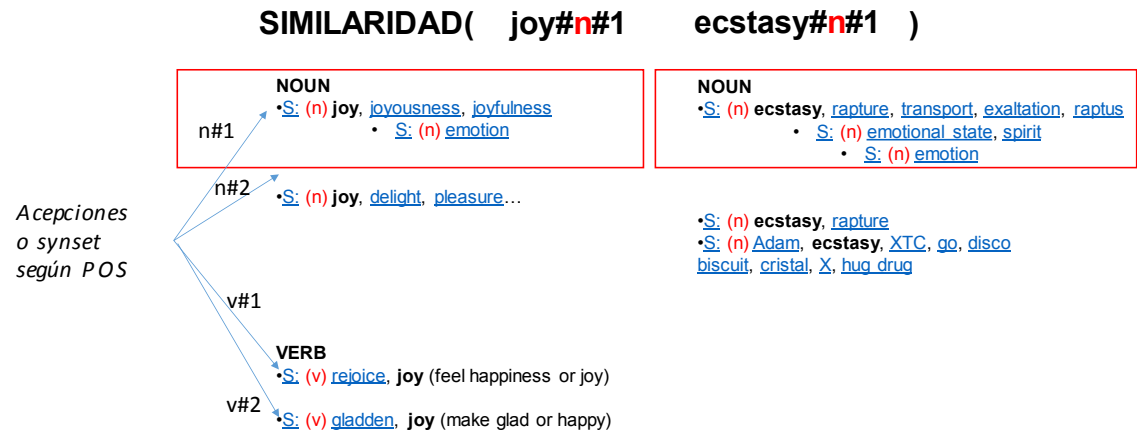


Figura 4-5: Ejemplo de similaridad entre palabras con un mismo POS.

### 4.3. Selección de Subconjunto de Análisis

Desde el proceso anterior, se obtiene un listado de 640 palabras que tienen raíz afectiva en *Wordnet*. Sin embargo, dado que la primera validación de los resultados, se realiza en comparación al ordenamiento manual, se consulta a un experto para seleccionar 10 palabras por cada clase afectiva. Con este listado, se realiza una encuesta de ordenamiento según la intensidad afectiva. En la Tabla 3 se detallan las palabras seleccionadas para cada clase afectiva.




Tabla 3. Términos seleccionados para la encuesta y etapa de validación

<b>Clases afectivas</b>							
<b>Anger</b>	<b>Anticipation</b>	<b>Disgust</b>	<b>Fear</b>	<b>Joy</b>	<b>Sadness</b>	<b>Trust</b>	<b>Surprise</b>
discontent	fancy	dislike	worry	amusement	unhappiness	liking	stab
upset	whim	distaste	alarm	gladness	pity	leaning	startle
resentment	aspiration	disgust	scare	delight	distress	respect	jolt
irritation	craving	displeasure	anxiety	cheerfulness	regret	approbation	scare
indignation	expectation	aversion	trepidation	happiness	melancholy	optimism	surprise
frustration	eagerness	revulsion	fright	joy	sorrow	confidence	excitement
hostility	excitement	hatred	fear	rejoicing	depression	loyalty	astonishment
wrath	ambition	horror	dread	glee	suffering	reverence	amaze
fury	optimism	animosity	panic	exhilaration	agony	hope	stunned
rage	hopefulness	loathing	terror	ecstasy	grief	admiration	shock


#### 4.4. Encuesta de opinión sobre intensidad afectiva

La encuesta para el ordenamiento de las palabras según la intensidad afectiva se distribuye a personas con estudios universitarios del idioma inglés o nativos. A través de la encuesta, el usuario puede seleccionar la o las clases afectivas que prefiere contestar. Para cada clase afectiva aparecen 10 palabras en un recuadro, que deben ser ordenadas de la menos intensa a la más intensa dentro de la clase afectiva, además se le solicita indicar el nivel de conocimiento en el idioma inglés, siendo los 4 niveles posibles Básico, Intermedio, Avanzado y Nativo (ver Figura 4-6). Finalmente, las respuestas obtenidas son validadas con un experto, con el objetivo de eliminar las que visiblemente fueron contestadas sin comprender las instrucciones entregadas.

Language: English Spanish



Affect Analisis Survey



UNIVERSIDAD DEL BÍO-BÍO

This survey is part of an investigation in the area of affective analysis of texts. The words consulted were extracted from a validated list in the field of psychology (affective lexicon, EMOLEX [1]). This lexicon is based on the premise that these words are used when expressing some affection such as anger, joy, sadness, anticipation, surprise, fear, trust, disgust [2].

1

Please select the affect class to answer the survey

Choose an option

2

Please indicate your level of English.

☐ Basic ☐ Intermediate ☐ Advanced ☐ Native

Below are 10 words that you should sort from the least intense to the most intense within the selected affective class.

3

To sort you must drag the words to the yellow rectangle on the right side

→

The number next to the word indicates the intensity: 1. Minimum intensity 10. Maximum intensity

4

Send Survey

[1] Mohammad, S.M., Turney, P.D.: Crowdsourcing a word-emotion association lexicon. In: Computational Intelligence. pp. 436–465 (2013).

[2] Plutchik, R.: The emotions. University Press of America (1991).

Figura 4-6. Encuesta de opinión. *screenshot* Aplicación Web

Los detalles de la implementación y los resultados obtenidos en la encuesta se entregan en el Anexo B, página 8-41, de este informe.

## 4.5. Modelo de intensidad afectiva

El modelo propuesto se basa en el uso de las métricas de similaridad implementadas en *Wordnet Similarity*, aplicadas al lexicón de afectos seleccionado [21–23], el cuál consta de 3 etapas principales.

- Etapa de homologación.
- Etapa de cálculo.
- Etapa de ordenamiento.

4-4

### 4.5.1. Etapa de homologación

En esta etapa se verifica que los términos que son extremos afectivos superiores de la clasificación afectiva de Plutchik, tengan una raíz afectiva dentro de *Wordnet-Affect*. En el caso que alguno de los extremos no cumpla el requisito se debe sustituir por alguna palabra que tenga un significado similar y que si cumpla con el requisito. La homologación es realizada en forma manual por un experto en idioma inglés. Por ejemplo, en este proceso se verifica que el término *vigilance*, extremo afectivo superior de la clase afectiva *Anticipation*, no tiene raíz afectiva, por lo que se reemplaza por el término *hopefulness*.

La verificación se realiza mediante una aplicación construida para este propósito.

### 4.5.2. Etapa de cálculo

En esta etapa se realiza el cálculo de similaridad entre cada una de las palabras y el extremo superior de la categoría afectiva a la que pertenecen, para cada par palabra-extremo se aplican las seis métricas de similaridad [11–16]. Adicionalmente, se verifican las restricciones que impone cada métrica de similaridad para poder ejecutar el cálculo (por ejemplo, relación POS: sustantivo-sustantivo, verbo-verbo). Esta etapa se realiza mediante una aplicación construida para este fin, basada en los módulos *wordnet* y *wordnet\_ic* de NLTK<sup>4</sup>, los cuáles proveen las funciones de recuperación de IC y cálculo de similaridad. Los detalles de la implementación se encuentran en el Anexo C, página 8-46, de este informe.

### 4.5.3. Etapa de ordenamiento

En esta etapa se recogen los cálculos realizados anteriormente, luego se separan por clase afectiva, y por cada clase se preparan los resultados de las 6 métricas aplicadas. Esto se realiza de forma automática, a través de una aplicación construida con este propósito.

---

<sup>4</sup> Natural Language Toolkit - <http://www.nltk.org/>

## 4.6. Correlación entre expertos v/s modelo

La encuesta fue contestada por 39 personas con estudios del idioma inglés. Luego de realizar la revisión con un experto, se considera que 4 encuestas no fueron contestadas siguiendo las instrucciones, por lo que son descartadas, dejando 35 respuestas válidas distribuidas entre las 8 clases afectivas consultadas, con un mínimo de 3 respuestas por cada clase afectiva.

La concordancia y porcentaje de acuerdo entre las respuestas se analizó utilizando la medida Kappa de Randolph [30], basada en la medida Kappa de Fleiss [31], que permite realizar un análisis descriptivo, obteniendo el grado de concordancia entre múltiples evaluadores al etiquetar sujetos según un conjunto de clases. Este cálculo se realizó en una aplicación en línea<sup>5</sup>.

Cada evaluador entrega el orden de las 10 palabras de la clase, desde menos a más intensas, por ejemplo, para las palabras de la clase *anger* un evaluador otorga el orden de las palabras.

	10 categorías	5 categorías
discontent	1	1
upset	2	
resentment	3	
irritation	4	2
indignation	5	
frustration	6	3
hostility	7	
wrath	8	4
fury	9	
rage	10	5

Figura 4-7: Categorías para el análisis de concordancia.

Si bien existen 10 categorías para el análisis de concordancia, es válido que éstas se reagrupen en 5 categorías para representar las posiciones en el orden

---

<sup>5</sup> Randolph, J. J. (2008). Online Kappa Calculator [Computer software]. <http://justus.randolph.name/kappa>

de intensidad, en este caso: {1-2}, {3-4}, {5-6}, {7-8} y {9-10}. Lo anterior, se basa en que “es igualmente bueno que un experto indique que la palabra *wrath* es la séptima u octava en intensidad”. En general, la percepción de la intensidad es una evaluación altamente subjetiva e incluso personal para cada experto, por lo que bien podría evaluarse difusamente. La concordancia utilizando la medida de Kappa se ve directamente afectada por la cantidad de categorías a evaluar.

En la Tabla 4 se visualizan las clases afectivas junto al porcentaje de acuerdo y el índice Kappa de las respuestas de los encuestados. Es importante destacar que, al analizar respuestas en torno a emociones, es presumible que se obtengan bajos niveles de concordancia. El detalle de las respuestas se encuentra en el Anexo D, página 8-53.

Tabla 4. Concordancia e índice kappa por clase afectiva

Clase afectiva	Nº de encuestas válidas	Porcentaje de acuerdo	Índice Kappa
Anger	6	43 %	0,29
Anticipation	3	23 %	0,04
Disgust	4	36 %	0,20
Fear	3	60 %	0,50
Joy	8	40 %	0,25
Sadness	5	34 %	0,17
Trust	3	36 %	0,20
Surprise	3	30 %	0,13

Posteriormente se analizan los resultados del ordenamiento entregado por el modelo versus el ordenamiento obtenido desde la encuesta, esto se realiza utilizando el coeficiente de correlación de Pearson [32], el que indica el nivel de correlación lineal entre 2 variables cuantitativas. Para este caso se considera como primera variable al ordenamiento promedio entregado por las respuestas de la encuesta, y como segunda variable al ordenamiento entregado por el modelo. Este análisis se realiza por cada métrica de similaridad revisada (path, wup, lch, res, jcn, lin) y en cada una de las categorías afectivas. En la Tabla 5 se visualizan los resultados del análisis.

Tabla 5. Correlación de Pearson entre resultados de la encuesta y resultados del modelo

Clase	path[11]	wup[12]	lch[13]	res[14]	jcn[15]	lin[16]
Anger	0,83	0,84	0,93	0,96	0,98	0,98
Anticipation	0,09	0,09	0,11	0,06	0,64	0,64
Disgust	0,87	0,87	0,87	0,71	0,73	0,67
Fear	0,78	0,78	0,81	0,85	0,87	0,89
Joy	0,70	0,70	0,65	0,65	0,64	0,61
Sadness	0,14	0,14	0,14	0,37	-0,006	-0,006
Trust	-0,27	-0,27	-0,27	-0,18	-0,24	-0,24
Surprise	0,36	0,36	0,36	0,70	0,36	0,35

## 4.7. Correlación entre [29] v/s modelo

En esta sección se presentan los resultados de la evaluación de un subconjunto de palabras extraídas de un lexicón sin intensidad afectiva, que ha sido etiquetado utilizando el modelo propuesto en este trabajo. Se utiliza el coeficiente de correlación de Pearson [32] para determinar el grado de efectividad del modelo que determina la intensidad y comprobar si existe alguna correlación con la intensidad de las palabras en [29].

El valor del índice de correlación Pearson varía en el intervalo  $[-1,1]$ , indicando el signo el sentido de la relación (r representa al índice Pearson)

- Si  $r = 1$ , existe una correlación positiva perfecta. El índice indica una dependencia total entre las dos variables denominada relación directa: cuando una de ellas aumenta, la otra también lo hace en proporción constante.
- Si  $0 < r < 1$ , existe una correlación positiva.
- Si  $r = 0$ , no existe relación lineal. Pero esto no necesariamente implica que las variables son independientes: pueden existir todavía relaciones no lineales entre las dos variables.
- Si  $-1 < r < 0$ , existe una correlación negativa.
- Si  $r = -1$ , existe una correlación negativa perfecta. El índice indica una dependencia total entre las dos variables llamada relación inversa: cuando una de ellas aumenta, la otra disminuye en proporción constante.

### 4.7.1. Corpus de Evaluación

El corpus de evaluación se extrae desde el lexicón etiquetado con intensidad afectiva propuesto en [29]. Se utiliza este lexicón ya que, al igual que

esta investigación, se basa en la clasificación afectiva de Plutchik, aunque cabe destacar que la propuesta utiliza un subconjunto de 4 clases, de las 8 clases base. Además utiliza como una de sus fuentes de datos el lexicón presentado en [23], el cual es el mismo que se utiliza como fuente para extraer las palabras a enriquecer por el modelo propuesto en esta investigación.

#### 4.7.2. Detalle de Resultados

A continuación, se presentan los resultados del análisis de correlación, en función del orden de los términos. Para el análisis se utilizan 10 términos para cada clase afectiva que son descritos en la Tabla 3. Además los términos seleccionados existen en el lexicón propuesto en [29], lo que permite un análisis completo de cada clase afectiva.

Esta sección se divide en 4 sub-secciones, que se refieren a cada una de las clasificaciones afectivas a analizar.

#### Resultados de Clase afectiva Anger

La Tabla 6 presenta los resultados del análisis de correlación, usando el índice de Pearson, de la intensidad afectiva, en función del orden, para cada una de las métricas de similaridad con las que fue aplicado el modelo. En la Figura 4-8 se presentan los resultados de forma gráfica. En esta clase afectiva, el índice de correlación más alto lo tiene la aplicación del modelo propuesto a través de la métrica de similaridad Resnik [14]. El menor índice de correlación para esta clase lo tiene la aplicación del modelo utilizando las métricas Jiang & Conrath[15] y la métrica Lin[16].

Tabla 6. Coeficiente de Pearson para Clase Anger. Lexicón vs Propuesta

	Orden de intensidad de cada palabra según la métrica de similaridad aplicada.						
Palabras	Lexicón [29]	path	wup	lch	res	jcn	lin
discontent	1	3	3	4	2	3	3
upset	3	1	2	3	1	1	1
resentment	6	4	4	1	4	4	4
irritation	2	6	6	6	6	7	7
indignation	5	7	7	7	7	6	6

Orden de intensidad de cada palabra según la métrica de similitud aplicada.							
Palabras	Lexicón [29]	path	wup	lch	res	jcn	lin
frustration	4	2	1	2	3	2	2
hostility	7	5	5	5	5	5	5
wrath	8	8	8	8	8	8	8
fury	10	9	9	9	9	9	9
rage	9	10	10	10	10	10	10
Coef. Pearson		0,75	0,73	0,61	0,78	0,71	0,71

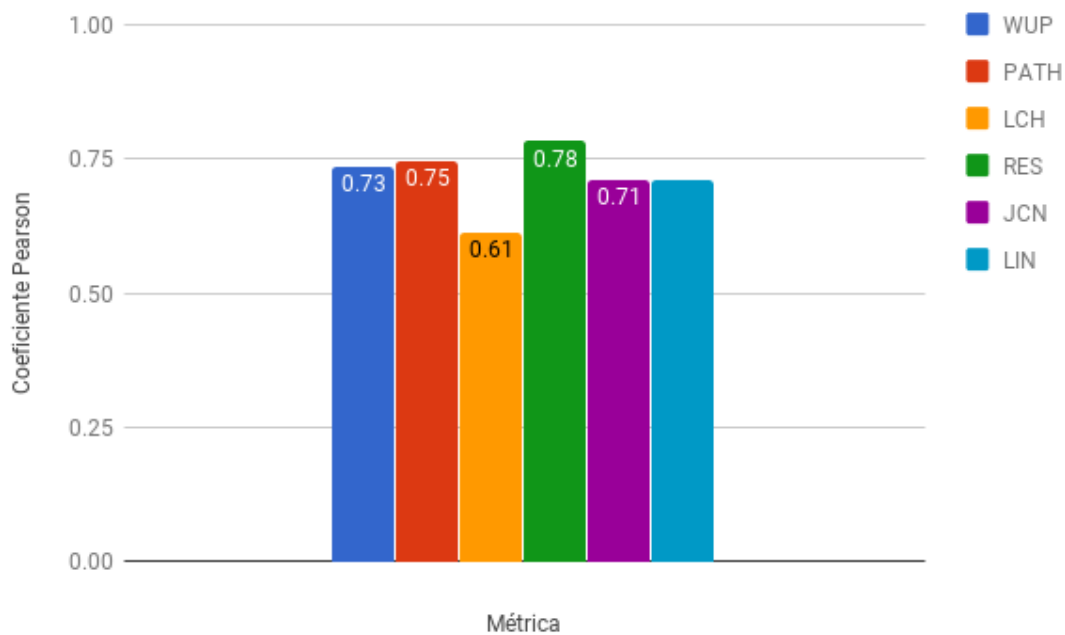


Figura 4-8. Coeficiente de Pearson con respecto a lexicón propuesto en [29]

## Resultados de clase afectiva Fear

La Tabla 7 presenta los resultados del análisis de correlación usando el índice de Pearson, en función del orden, para cada una de las métricas de similitud con las que fue aplicado el modelo. En la Figura 4-9 se presentan los resultados de forma gráfica. En esta clase afectiva el coeficiente de correlación más alto es nuevamente para la métrica Resnik [14] aplicada al modelo, mientras que la menor correlación es compartida por varias métricas. Sin embargo, las diferencias en la correlación son mínimas entre las métricas.



Tabla 7. Coeficiente de Pearson para Clase Fear. Lexicón vs Propuesta

Palabras	Lexicón [29]	path	wup	lch	res	jcn	lin
worry	1	1	1	1	1	1	1
alarm	3	4	4	4	3	4	4
scare	7	5	5	5	4	8	8
anxiety	2	3	3	2	2	3	2
trepidation	4	2	2	3	5	2	3
fright	8	7	7	7	6	6	6
fear	6	8	8	8	7	7	7
dread	9	6	6	6	8	5	5
panic	5	9	9	9	9	9	9
terror	10	10	10	10	10	10	10
Coef. Pearson		0,76	0,76	0,78	0,81	0,73	0,76

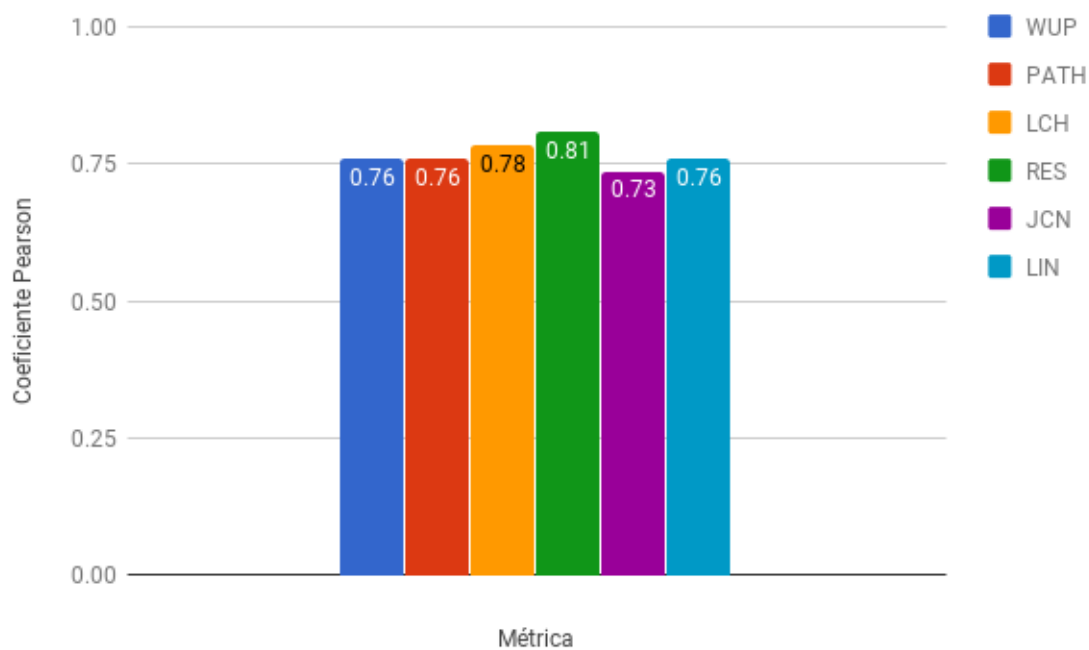


Figura 4-9. Coeficiente de Pearson para la clase afectiva *Fear*. Con respecto al lexicón creado en [29]

## Resultados de clase afectiva Sadness

La Tabla 8 presenta los resultados del análisis de correlación usando el índice de Pearson, de la intensidad afectiva, en función del orden, para cada una de las métricas de similaridad con las que fue aplicado el modelo. En la Figura 4-10 se presentan los resultados de forma gráfica. En esta clase afectiva, el índice de correlación es bajo en general, siendo el más destacable el entregado por la aplicación del modelo propuesto a través de la métrica de similaridad Resnik [14]. El menor índice de correlación para esta clase lo tiene la aplicación del modelo utilizando las métricas Jiang & Conrath[15] y la métrica Lin[16].

Tabla 8. Coeficiente de Pearson para Clase *Sadness*. Lexicón vs Propuesta

Palabras	Lexicón [29]	path	wup	lch	res	jcn	lin
unhappiness	5	2	2	2	1	3	3
pity	1	4	4	4	2	4	4
distress	2	5	5	5	3	5	5
regret	4	7	7	7	7	8	8
melancholy	3	8	8	8	8	7	7
sorrow	6	1	1	1	4	2	2
depression	10	9	9	9	9	9	9
suffering	7	6	6	6	5	6	6
agony	9	3	3	3	6	1	1
grief	8	10	10	10	10	10	10
Coef. Pearson		0,22	0,22	0,22	0,55	0,15	0,15

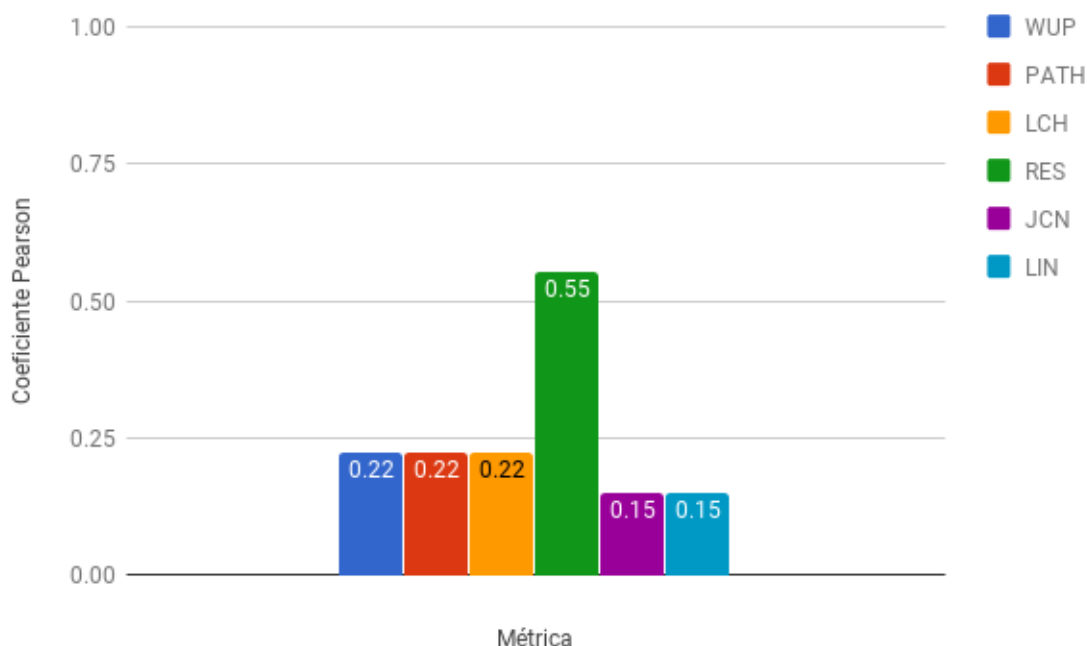


Figura 4-10. Coeficiente de Pearson para la clase afectiva *Sadness*. Con respecto al lexicón creado en [29]

## Resultados de clase afectiva Joy

La Tabla 9 presenta los resultados del análisis de correlación usando el índice de Pearson, en función del orden, para cada una de las métricas de similaridad con las que fue aplicado el modelo. En la Figura 4-11 se presentan los resultados de forma gráfica. En esta clase afectiva el coeficiente de correlación más alto es compartido por las métricas Path [11] y Wu&Palmer [12], aplicadas al modelo, mientras que la menor correlación es entregada por la aplicación de la métrica Jiang&Conrath [15]. Sin embargo, las diferencias en la correlación son mínimas entre las métricas.

Tabla 9. Coeficiente de Pearson para Clase *Joy*. Lexicón vs Propuesta.

Palabras	Lexicón [29]	path	wup	lch	res	jcn	lin
amusement	1	1	1	2	2	5	5
gladness	4	3	3	3	3	1	1
delight	3	4	4	4	4	6	6

Palabras	Lexicón [29]	path	wup	lch	res	jcن	lin
cheerfulness	2	5	5	5	5	2	2
happiness	10	9	9	9	9	8	9
joy	9	8	8	8	8	9	8
rejoicing	5	6	6	6	6	3	3
glee	6	2	2	1	1	4	4
exhilaration	7	7	7	7	7	7	7
ecstasy	8	10	10	10	10	10	10
Coef. Pearson		0,79	0,79	0,73	0,73	0,70	0,71

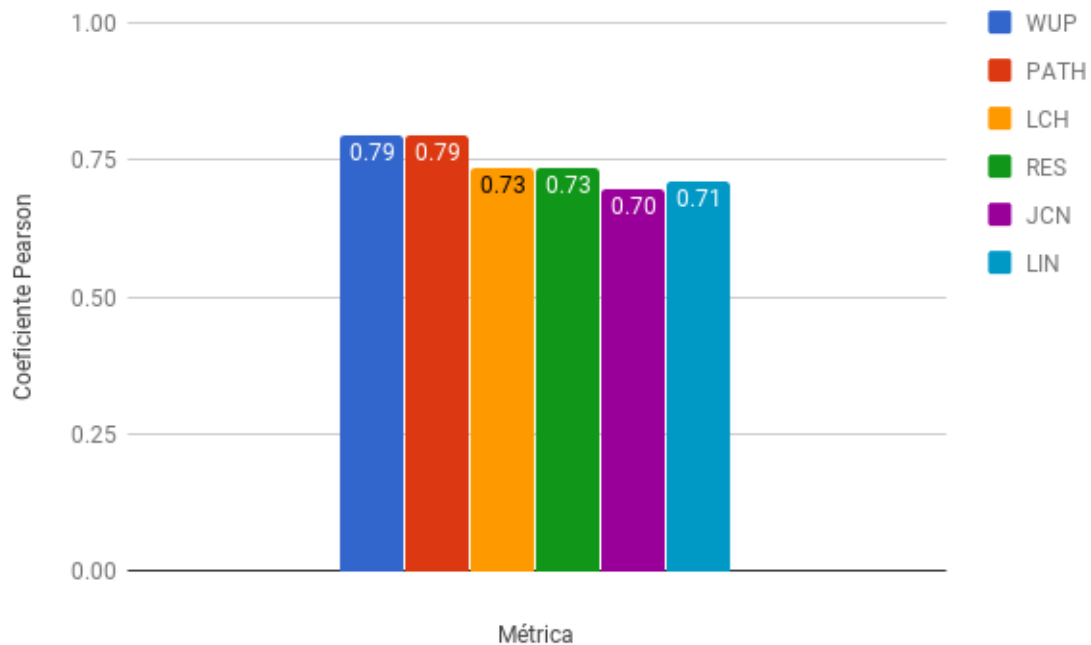


Figura 4-11. Coeficiente de Pearson para la clase afectiva Joy, con respecto al lexicón creado en [29]

## 4.8. Conclusión del capítulo

Se presenta un modelo para calcular la intensidad afectiva de una palabra en una clase. El modelo se basa en la clasificación afectiva de Plutchik y se prueba utilizando un lexicón construido en [23]. Tal como se explicó antes, dicho

recurso fue seleccionado ya que utiliza *Wordnet-Affect* como fuente de datos y utiliza la clasificación afectiva de Plutchik para la organización de las palabras.

Se construye una encuesta donde 35 evaluadores construyen una medida de intensidad, en función del orden de los términos, para un subconjunto de 10 palabras de cada clase afectiva, donde la ubicación de la palabra indica la intensidad afectiva dentro de la clase. Una ubicación en puestos más bajos indica menor intensidad, mientras una ubicación más alta indica mayor intensidad. Por último, se evalúa la correlación de la medida de intensidad entregada por el modelo propuesto, en función del orden de los términos, con el orden entregado por los evaluadores.

Se realiza el análisis de correlación mediante el índice Pearson, entre el corpus de evaluación extraído del lexicón propuesto en [29] y los resultados de la aplicación del modelo con cada una de las métricas de intensidad para cada clase afectiva a analizar.

De las 6 métricas de intensidad utilizadas para evaluar el modelo, 3 se basan en estructura [11–13] y 3 en contenido informativo (IC) [14–16]. Al calcular el índice de correlación promedio por tipo de métrica, se obtiene un valor de 0,62 para las 2 agrupaciones.

Al evaluar el modelo con las 6 métricas de similaridad analizadas, destaca Resnik [14] con un índice de correlación promedio de 0,72 y un mínimo de 0,55.

A continuación, se presentan los resultados generales, que entregan comparaciones entre el rendimiento del modelo por cada métrica.

#### **4.8.1. Resultados generales del modelo de intensidad afectiva.**

Al comparar los resultados por cada clase afectiva estudiada, y por cada métrica de similaridad aplicada al modelo de intensidad afectiva propuesto, se visualiza que en la clase afectiva *Anger*, se obtienen un índice de correlación mínimo de 0,61 y un índice máximo de 0,78, con un índice promedio de 0,72. Para la clase afectiva *Fear* se obtiene un índice de correlación máximo de 0,81 y un índice mínimo de 0,73, con un promedio de 0,77.

En la clase afectiva *Sadness* se obtienen resultados más dispersos. El índice de correlación máximo de 0,55 y un índice mínimo de 0,15, con un promedio de 0,25.

Por último, en la clase afectiva *Joy* se obtiene un índice de correlación máximo de 0,79 y un índice mínimo de 0,70, con un índice promedio de 0,74.

En la Tabla 10 se presentan los índices de correlación de cada métrica en cada clase afectiva, en comparación a los resultados del corpus. En la Figura 4-12 se visualiza la misma información en forma gráfica, y en la Figura 4-13 se visualiza la correlación promedio de cada métrica.

Por último en la Tabla 11 se presenta la correlación entre los resultados de la encuesta realizada para esta investigación, en comparación al corpus extraído del lexicon propuesto en [29].

A continuación, en el próximo capítulo se revisarán temas de discusión relacionados al trabajo realizado y a la luz de los resultados expuestos en este capítulo.

Tabla 10. Índice de Pearson por cada métrica de similaridad para cada clase afectiva.

	path	wup	lch	res	jcn	lin
Anger	0,73	0,75	0,61	0,78	0,71	0,71
Fear	0,76	0,76	0,78	0,81	0,73	0,76
Sadness	0,22	0,22	0,22	0,55	0,15	0,15
Joy	0,79	0,79	0,73	0,73	0,70	0,71

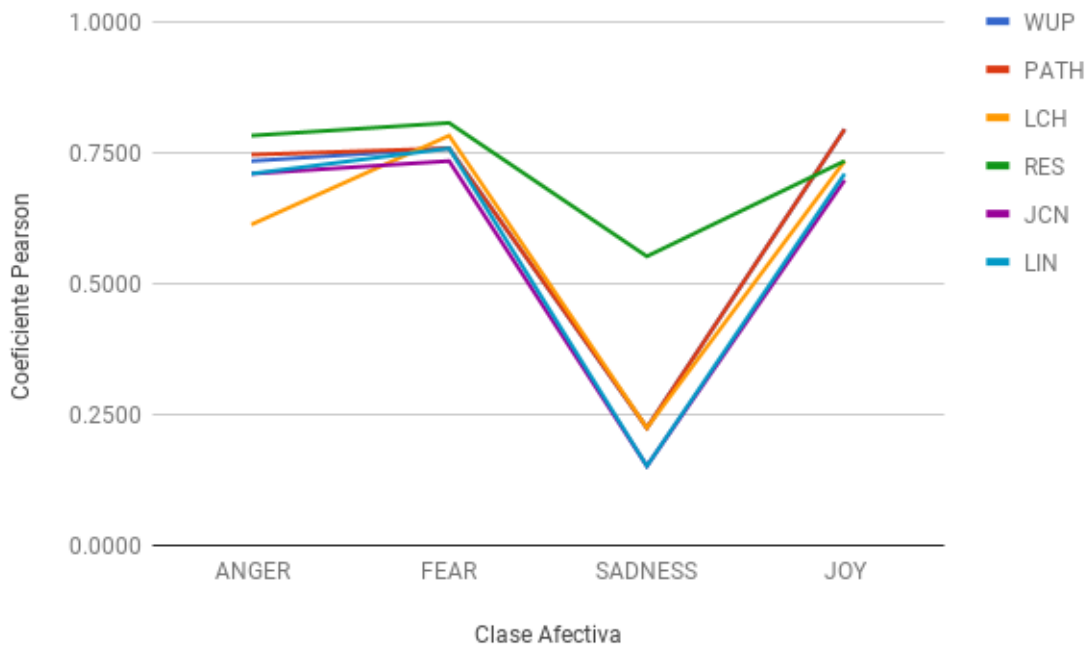


Figura 4-12. Índice de Pearson por cada métrica de similaridad para cada clase afectiva.

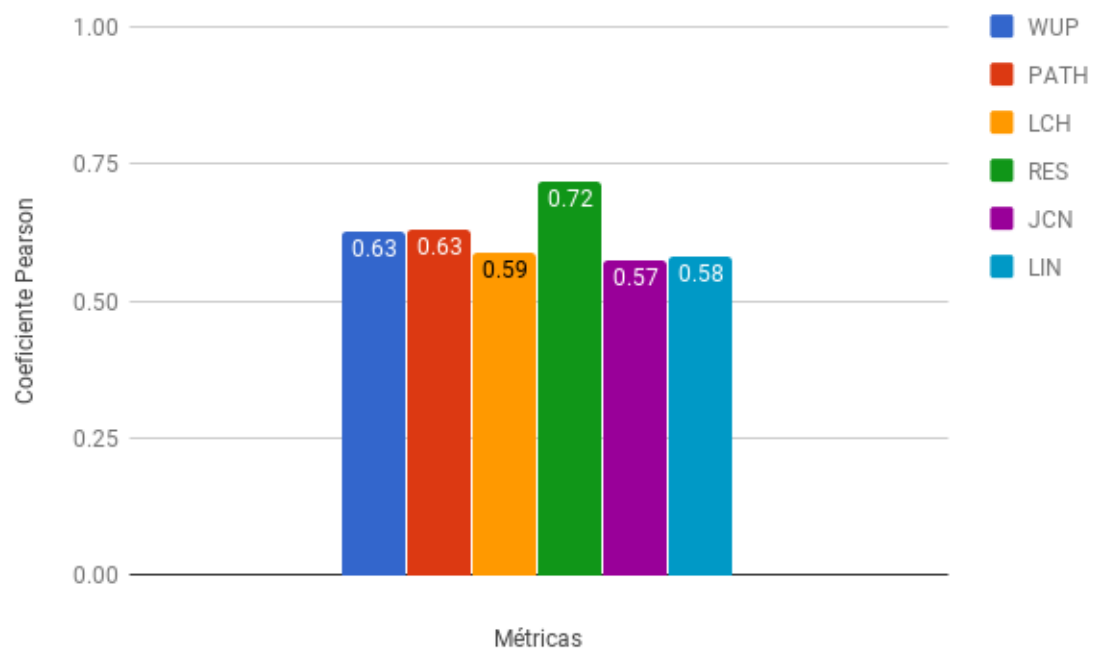


Figura 4-13. Índice de correlación en promedio para cada métrica de Similitud en las 4 clases afectivas revisadas.

Tabla 11. Índice Pearson de la encuesta en comparación al corpus de evaluación.

	Anger	Fear	Sadness	Joy
Encuesta	0,93	0,76	0,84	0,65

## Capítulo 5: Evaluación: Análisis de afectos utilizando un lexicón enriquecido con intensidad

En el capítulo anterior se propone un modelo que permite agregar, en forma automática, la intensidad afectiva a palabras que tengan al menos un ancestro afectivo en *Wordnet*. El modelo se evalúa por medio de un análisis de correlación con un subconjunto de 10 palabras por cada clase afectiva, extraídas del lexicón EmoLex, y enriquecidas a través del modelo, versus la intensidad del mismo conjunto de palabras, extraído del lexicón propuesto en [29].

En este capítulo se utiliza el modelo de intensidad automática propuesto, para enriquecer el listado de 640 palabras que cumplen las condiciones de filtrado descritas en el capítulo anterior. Con el listado enriquecido se diseña y ejecuta un experimento de evaluación.

En la primera sección de este capítulo se presenta el diseño del experimento, donde se detalla en qué consiste el mismo, junto a la forma de evaluar el rendimiento del lexicón enriquecido. En la segunda sección se detalla la ejecución del experimento, para luego, en la tercera sección, presentar los resultados de la evaluación del lexicón enriquecido con intensidad afectiva, creado en base al modelo propuesto en este trabajo.

### 5.1. Diseño del experimento

En esta sección se describe el diseño del experimento de evaluación del lexicón enriquecido con intensidad afectiva, en forma automática, construido en este trabajo.

A través de la aplicación del modelo de intensidad afectiva automática propuesto en esta tesis, se añade la medida de intensidad a las 640 palabras que se obtienen luego de aplicar los filtros descritos en el capítulo anterior. Cabe destacar que el modelo de intensidad se aplica con 6 métricas de similaridad, por lo que para cada palabra se pueden obtener hasta 6 medidas de intensidad diferentes.



Para obtener el corpus que permita evaluar el lexicon enriquecido se realiza una búsqueda en los sitios web de investigadores que publican trabajos en el área del análisis de afecto basado en lexicon. De esta búsqueda se obtiene un corpus de *tweets* (publicaciones en la red social Twitter) etiquetado para las 4 clases afectivas estudiadas en la propuesta [29] y creado por el mismo autor del estudio, el cuál es referente en el área del AA, con numerosos artículos del área y material de análisis disponible en su sitio web<sup>6</sup>. Un extracto del corpus se puede ver en la Tabla 12.

Tabla 12. Extracto de corpus de tweets etiquetados con afecto

<b>Tweet</b>	<b>Clase afectiva</b>
My mind is raging and i just want to end it all	Anger
@leesyatt you are a cruel, cruel man. #therewillbeblood	Anger
First day of college feeling nervous	Fear
I'm just still . So happy .A blast	Joy
Vals always smiling I love it #gbbo	Joy
Going home is depressing	Sadness

Para realizar la evaluación del lexicon enriquecido se utiliza *Recall*, una métrica que es igual a la porción de documentos de una clase que son clasificados correctamente, en función de los casos correcta e incorrectamente clasificados. Para ellos considera los siguientes casos:

- True Positives (*tp*): elementos a los que el clasificador asignó la clase relevante y esta era correcta.
- False Positives (*fp*): elementos a los que el clasificador asignó la clase relevante y esta no era correcta.
- False Negatives (*fn*): elementos a los que el clasificador asignó la clase no relevante y esta no era correcta.
- True Negatives (*tn*): elementos a los que el clasificador asignó la clase no relevante y esta era correcta.

Esto se representa en la Tabla 13.

---

<sup>6</sup> <http://saifmohammad.com/WebPages/EmotionIntensity-SharedTask.html>

Tabla 13. Evaluación de los clasificadores

	Relevantes	No Relevantes
Seleccionados	tp	fp
No Seleccionados	fn	tn

En este caso, el clasificador es el lexicón enriquecido y se considera que la clase correcta es la clase indicada en el corpus utilizado.

Para cada clase afectiva se calcula *Recall* para la técnica basada en densidad de palabras, y para la técnica basada en intensidad de palabras, llamadas *Recall<sub>WD</sub>*, y *Recall<sub>WI</sub>* para objeto del análisis.

## 5.2. Experimento de Evaluación

El experimento de evaluación consiste en analizar, utilizando el lexicón enriquecido, cada frase del corpus de tweets para obtener la siguiente información:

- Clase afectiva obtenida desde el análisis por densidad de palabras por clase.
- Clase afectiva predominante obtenida desde el análisis por intensidad afectiva de las palabras.
- Intensidad afectiva de la frase en las 4 clases (*Anger*, *Fear*, *Sadness*, *Joy*)

Es importante recordar que el lexicón fue enriquecido a través del modelo de intensidad afectiva, el cual se aplica utilizando 6 métricas de similaridad, por lo que el lexicón incluye 6 valores de intensidad para cada palabra. Esto significa que en el análisis se entregan 6 resultados del análisis afectivo por frase.

El experimento se realiza a través de una aplicación construida para este propósito específico y disponible para su consulta<sup>7</sup>. La Aplicación consta de 2 secciones, la primera, que se visualiza en la Figura 5-1, es la sección de configuración, donde se incluyen los archivos que se refieren al lexicón para cada clase afectiva y cada métrica de similaridad con la que fue aplicado el modelo.

---

<sup>7</sup> <http://somos.ubiobio.cl/analisis/>

Además, permite indicar el texto que se desea analizar, el texto donde están los *stopwords* (palabras que no son incluidas en el análisis) que se desean utilizar, y el porcentaje mínimo de palabras de la frase que deben estar incluidas en el lexicon para que se realice el análisis afectivo.

Los detalles de la implementación de la herramienta desarrollada se encuentran en el Anexo C, página 8-48, de este informe.

## Análisis Afectivo IntensidadAutomática

Esta aplicación permite analizar afectivamente un listado de textos

Configuración

Resultados

Configuración

Antes de iniciar el análisis debe subir los archivos que se solicitan, incluyendo lexicon etiquetado con intensidad afectiva y textos a analizar  
Formato lexicon: [palabra],[tipo],[pos],[intensidad]  
Formato textos: [cod]@[frase]@[emocion]@[intensidad]

Se ha completado el Análisis afectivo

Métrica	Lexicón Anger	Lexicón Fear	Lexicón Sadness	Lexicón Joy
PATH	<div>Seleccionar archivo</div> angerPath.csv	<div>Seleccionar archivo</div> fearPath.csv	<div>Seleccionar archivo</div> sadnessPath.csv	<div>Seleccionar archivo</div> joyPath.csv
WUP	<div>Seleccionar archivo</div> angerWUP.csv	<div>Seleccionar archivo</div> fearWUP.csv	<div>Seleccionar archivo</div> sadnessWUP.csv	<div>Seleccionar archivo</div> joyWUP.csv
LCH	<div>Seleccionar archivo</div> angerLCH.csv	<div>Seleccionar archivo</div> fearLCH.csv	<div>Seleccionar archivo</div> sadnessLCH.csv	<div>Seleccionar archivo</div> joyLCH.csv
RES	<div>Seleccionar archivo</div> angerRES.csv	<div>Seleccionar archivo</div> fearRES.csv	<div>Seleccionar archivo</div> sadnessRES.csv	<div>Seleccionar archivo</div> joyRES.csv
JCN	<div>Seleccionar archivo</div> angerJCN.csv	<div>Seleccionar archivo</div> fearJCN.csv	<div>Seleccionar archivo</div> sadnessJCN.csv	<div>Seleccionar archivo</div> joyJCN.csv
LIN	<div>Seleccionar archivo</div> angerLIN.csv	<div>Seleccionar archivo</div> fearLIN.csv	<div>Seleccionar archivo</div> sadnessLIN.csv	<div>Seleccionar archivo</div> joyLIN.csv

Texto a analizar

Seleccionar archivo

 textos-fear.txt

Stopwords

Seleccionar archivo

 stopwords.txt

Cobertura Mínima

34 %

Comenzar Análisis

Figura 5-1. Aplicación para análisis afectivo – Configuración

En la segunda sección, que puede ser visualizada en la Figura 5-2, se presentan los resultados del análisis del texto a analizar, basado en los lexicones incluidos, el listado de *stopwords* agregado y el porcentaje mínimo de palabras de la frase que están contenidas en el lexicon.

## Análisis Afectivo IntensidadAutomática

Esta aplicación permite analizar afectivamente un listado de textos

Configuración

Resultados

Resultados del Análisis

Legenda:

Anger Fear Sadness Joy

Nº	Frase	Afecto - Intensidad	PATH	WUP	LCH	RES	JCN	LIN
0	I feel like I am drowning. #depression #anxiety #failure #worthless	fear - 0.979	Afecto: Sadness 0.909 0.909 0.909 Cobertura: 40.0%	Afecto: Sadness 0.909 0.909 0.909 Cobertura: 40.0%	Afecto: Sadness 0.921 0.921 0.909 Cobertura: 40.0%	Afecto: Sadness 0.909 0.909 0.909 Cobertura: 40.0%	Afecto: Sadness 0.909 0.909 0.909 Cobertura: 40.0%	Afecto: Sadness 0.909 0.909 0.909 Cobertura: 40.0%
2	I lost my blinders .... #panic	fear - 0.975	Afecto: Fear 0 0.909 0 0 Cobertura: 16.7%	Afecto: Fear 0 0.909 0 0 Cobertura: 16.7%	Afecto: Fear 0 0.909 0 0 Cobertura: 16.7%	Afecto: Fear 0 0.909 0 0 Cobertura: 16.7%	Afecto: Fear 0 0.909 0 0 Cobertura: 16.7%	Afecto: Fear 0 0.909 0 0 Cobertura: 16.7%
3	I feel like I am drowning. #depression #failure #worthless	fear - 0.938	Afecto: Sadness 0 0 0.909 0 Cobertura: 10.0%	Afecto: Sadness 0 0 0.909 0 Cobertura: 10.0%	Afecto: Sadness 0 0 0.909 0 Cobertura: 10.0%	Afecto: Sadness 0 0 0.909 0 Cobertura: 10.0%	Afecto: Sadness 0 0 0.909 0 Cobertura: 10.0%	Afecto: Sadness 0 0 0.909 0 Cobertura: 10.0%
4	This is the scariest American Horror Story out of all of them... I'm gonna have to watch in the daytime. #rightened	fear - 0.938	Afecto: Fear 0.909 0.909 0.909 Cobertura: 14.3%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 14.3%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 14.3%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 14.3%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 14.3%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 14.3%
5	@mgcartwork: I nearly started crying and having a full on panic attack after tatino bc of the crowds so I feel him	fear - 0.938	Afecto: Fear 0 0.909 0 0 Cobertura: 4.5%	Afecto: Fear 0 0.909 0 0 Cobertura: 4.5%	Afecto: Fear 0 0.909 0 0 Cobertura: 4.5%	Afecto: Fear 0 0.909 0 0 Cobertura: 4.5%	Afecto: Fear 0 0.909 0 0 Cobertura: 4.5%	Afecto: Fear 0 0.909 0 0 Cobertura: 4.5%
6	I have to finally tell my therapist about my sexuality ... last frontier ... not sure I can do it in the AM #fear #SingleGirlProblems	fear - 0.938	Afecto: Fear 0.909 0.909 0.909 Cobertura: 8.0%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 8.0%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 8.0%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 8.0%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 8.0%	Afecto: Fear 0.909 0.909 0.909 Cobertura: 8.0%

### 5.3. Detalle de Resultados de evaluación de lexicón enriquecido

A continuación, se presentan los resultados del análisis de rendimiento del lexicón enriquecido en forma automática. Se destaca que los resultados presentados se refieren a las frases del corpus de tweets, para las que se encuentra más de  $1/3$  ( $>33\%$ ) de las palabras en el lexicón enriquecido. Ya que el lexicón enriquecido contiene 640 palabras y representa el  $7,7\%$  de EmoLex, es predecible que pocas frases cumplan con este criterio. Sin embargo, es necesario para evaluar el comportamiento del lexicón en condiciones donde pueda realizar el análisis.

Para la clase afectiva *Anger* se analizan 856 tweets, de los cuales 11 cumplen con el criterio mínimo. Se destaca el análisis realizado con el lexicón generado con la métrica Lin[16], el que identifica las 11 frases dentro de la clase *Anger*. Los detalles de la clase *Anger* se visualizan en la Tabla 14.

Tabla 14. Métrica Recall para evaluación lexicón - Clase Anger

	Recall <sub>wd</sub>	Recall <sub>wl</sub>
PATH	0,36	0,82
WUP	0,36	0,82
LCH	0,36	0,82
RES	0,36	0,82
JCN	0,36	0,82
LIN	0,36	1,00

Dentro de la clase afectiva *Fear* se analizan con 1145 tweets, de los cuales 25 cumplen con el criterio mínimo. Se destaca la gran diferencia al utilizar la densidad de palabras, versus la utilización de la intensidad como medio de identificación del afecto predominante. Los detalles de la clase *Fear* se visualizan en la Tabla 15.

Tabla 15. Métrica Recall para evaluación lexicón - Clase Fear

	<b>Recall<sub>WD</sub></b>	<b>Recall<sub>WI</sub></b>
PATH	0,24	0,63
WUP	0,24	0,63
LCH	0,16	0,60
RES	0,16	0,63
JCN	0,16	0,60
LIN	0,16	0,20

En el caso de la clase afectiva *Sadness* se analizan 786 tweets, de los cuales 12 cumplen con el criterio mínimo. En esta clase destaca el hecho de que no existen mejoras al realizar el análisis utilizando la intensidad afectiva. Los detalles de la clase *Sadness* se visualizan en la Tabla 16.

Tabla 16. Métrica Recall para evaluación lexicón - Clase Sadness

	<b>Recall<sub>WD</sub></b>	<b>Recall<sub>WI</sub></b>
PATH	0,29	0,29
WUP	0,29	0,29
LCH	0,29	0,29
RES	0,29	0,29
JCN	0,29	0,29
LIN	0,29	0,29

En el caso de la clase afectiva *Joy* se analizan 822 tweets, de los cuales sólo 3 cumplen con el criterio mínimo. Por esta razón se rebaja el criterio de inclusión a un 30% de cobertura del lexicón dentro de la frase. Con esta rebaja se obtienen 6 tweets que cumplen el criterio. En esta clase destaca el hecho de que, aunque no existen mejoras al realizar el análisis utilizando la intensidad afectiva, los valores de Recall para el análisis basado en densidad de palabras son bastante altos. Los detalles de la clase *Joy* se visualizan en la tabla 17.

Tabla 17. Métrica Recall para evaluación lexicón - Clase Joy

	<b>Recall<sub>WD</sub></b>	<b>Recall<sub>WI</sub></b>
PATH	0,83	0,83
WUP	0,83	0,83
LCH	0,83	0,83

RES	0,83	0,83
JCN	0,83	0,83
LIN	0,83	0,83

## 5.4. Conclusión del capítulo

En este capítulo se describe el experimento destinado a evaluar el rendimiento del lexicón de afectos enriquecido mediante el modelo de intensidad automática propuesto en esta tesis. En términos generales, al analizar el rendimiento del lexicón cuando es evaluado desde la densidad de los conceptos es, en promedio de un 42%, y al evaluar las frases utilizando la intensidad afectiva, obtiene un rendimiento promedio de 62%. Si se realiza el mismo ejercicio, pero detallado por clase afectiva se obtiene un rendimiento promedio de 36% para la clase *Anger*, al realizar el análisis por densidad de palabras, y un promedio de 85% al evaluar las frases utilizando la intensidad etiquetada. La clase afectiva *Fear* entrega un rendimiento promedio de 19% al evaluar las frases en función de la densidad de las palabras, y un 55% de rendimiento promedio al utilizar la intensidad afectiva en el análisis. La evaluación de la clase afectiva *Sadness* es la que entrega los peores resultados, con un rendimiento promedio de 29%, sin que la forma de análisis afecte los resultados. Por último, el rendimiento promedio del lexicón enriquecido en la clase Joy entrega un rendimiento promedio de 83%, sin que se registren diferencias al cambiar el tipo de análisis realizado. La Figura 5-3 representa esta información gráficamente, mientras que la Figura 5-4 muestra el rendimiento detallado por cada una de las clases afectivas, métrica de similaridad utilizada, y método de análisis (densidad de palabras o intensidad de las palabras).

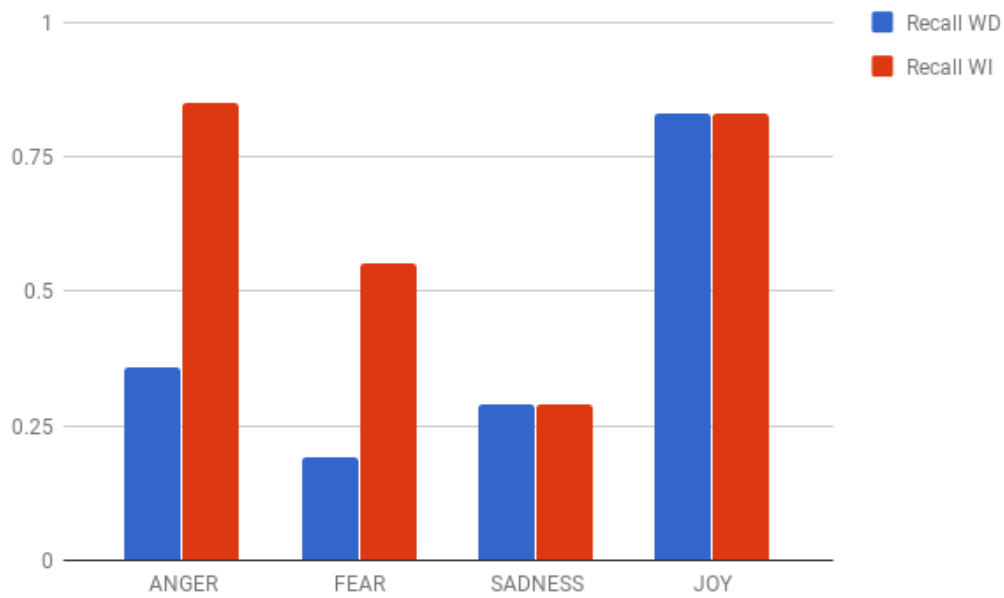


Figura 5-3. Comparativa entre método de análisis por clase afectiva

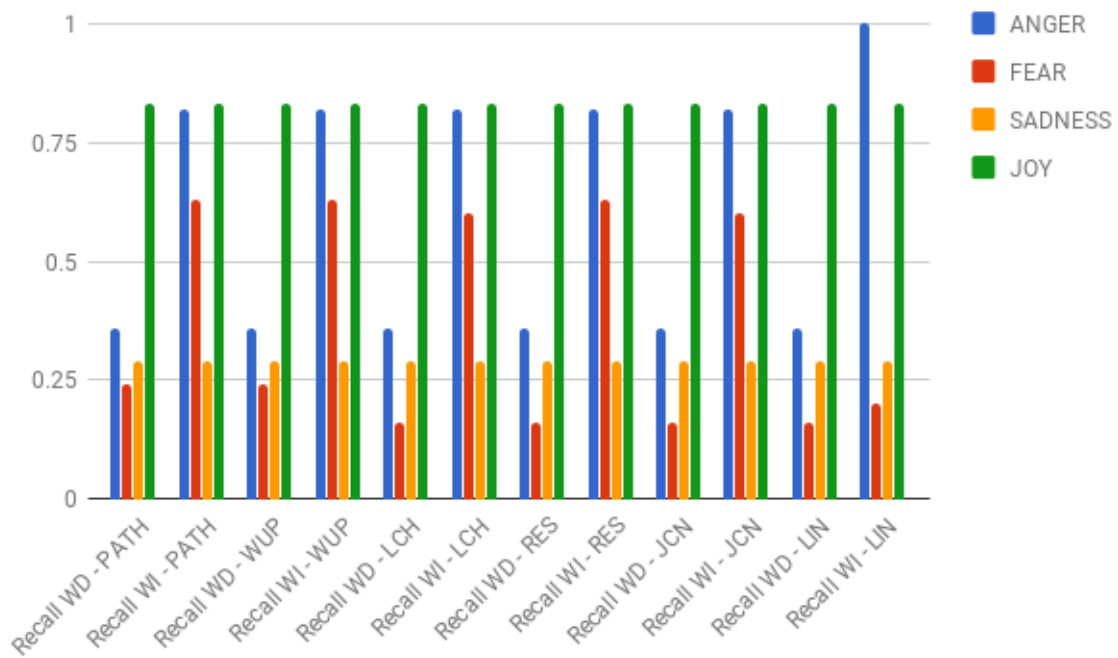


Figura 5-4. Comparativa de Rendimiento de cada métrica y tipo de análisis por clase afectiva

## Capítulo 6: Discusión

A continuación, se presenta la discusión de los resultados de la evaluación del modelo de intensidad automática propuesto y del experimento realizado para evaluar el rendimiento del lexicón enriquecido con intensidad afectiva.

### 6.1. Discusión de los resultados del modelo de intensidad afectiva automática

En términos generales, y tomando como referencia que se busca un índice de correlación lo más cercano a 1 como ideal, el modelo propuesto tiene un rendimiento promedio de un 62% tomando en cuenta todas las métricas de similaridad revisadas. El resultado es deficiente en comparación a otros trabajos revisados en esta investigación, pero es aceptable considerando la escasa o nula existencia de trabajos que indiquen la intensidad afectiva en forma automática.

Es interesante el análisis que surge a la luz de los resultados del experimento, ya que se obtienen rendimientos bastantes dispares dependiendo de la clase afectiva. Para las clases *Anger*, *Fear* y *Joy*, el modelo de intensidad afectiva se comporta de manera similar en cuanto al índice de correlación, sin importar la métrica de similaridad utilizada. Sin embargo, al evaluar el modelo en la clase afectiva *Sadness* los índices de correlación bajan hasta el 0,15, exceptuando a la métrica Resnik, con la que se logra un índice de 0,55, que se puede considerar aceptable.

Por otra parte, al realizar el análisis desde el punto de vista de las métricas revisadas, en las que están basadas en estructura [11–13] se observa un rendimiento promedio de 62%, y una variación máxima del índice de 0,12. Por parte de las métricas basadas en IC (Contenido Informativo)[14–16], se observa el mismo rendimiento promedio de 62%, sin embargo la variación máxima del índice es de 0,4, una variación alta para métricas que se basan en el mismo concepto base. Es importante destacar que el IC se genera a partir de un corpus, por lo que los resultados de las métricas basadas en IC pueden variar en función de la fuente de Contenido Informativo que se utilice.



En definitiva, el modelo propuesto tiene un rendimiento promedio de un 62%, con rendimientos aceptables en las clases afectivas *Anger*, *Fear* y *Joy*, y rendimientos deficientes en la clase afectiva *Sadness*.

La clase afectiva *Sadness* es la que obtiene los resultados más deficientes en cuanto al índice de correlación que obtiene, en comparación a las otras clases afectivas estudiadas, se cree que esto puede tener relación con la forma en que fue construido Wordnet y en particular *Wordnet-Affect* y con la complejidad que significa comprender la intensidad afectiva en emociones más complejas, esto ya que *Anger*, *Fear* y *Joy* son emociones primarias y más fáciles de comprender.

Referente al modelo de intensidad afectiva propuesto en este trabajo, se cree que es posible mejorar la estrategia para obtener el término que funciona como extremo superior de cada clase afectiva, y que permite calcular la intensidad afectiva usando las métricas de similaridad. Además, para este modelo se utilizó solamente el extremo superior de cada clase afectiva. Quizás se pueden obtener resultados distintos cambiando el modelo de intensidad afectiva, por ejemplo, utilizando los extremos superior e inferior para realizar el cálculo.

## 6.2. Discusión de los resultados del experimento de evaluación del lexicon enriquecido

En términos generales, según los resultados del experimento de evaluación del lexicon enriquecido, este tiene un rendimiento promedio de un 52% tomando en cuenta el análisis mediante la intensidad de todas las métricas incluidas en el lexicon y las dos formas de análisis, por densidad de palabras y por intensidad de palabras. Si se realiza el mismo ejercicio, pero esta vez haciendo la distinción entre cada forma de análisis, se obtiene un rendimiento de 42% si se realiza el análisis sólo por densidad de palabras, y un rendimiento de 62% si se realiza el análisis incluyendo la intensidad afectiva que fue añadida en esta investigación. Esto quiere decir que el análisis del lexicon enriquecido con intensidad afectiva fue un 20% más preciso que el análisis realizado con el mismo lexicon, pero sin intensidad afectiva. Este resultado se considera aceptable considerando la escasa existencia de trabajos que propongan estrategias para enriquecer lexicones, indicando la intensidad afectiva en forma automática.

A la luz de los resultados, es interesante realizar un análisis, ya que se obtienen rendimientos bastantes dispares dependiendo de la clase afectiva a la

que pertenecen las frases que se analizan. Para las clases *Anger* y *Fear*, el rendimiento mejora considerablemente al agregar el análisis por intensidad afectiva. Para la clase *Joy*, el rendimiento es bueno, y no hay cambios al agregar la intensidad al proceso de análisis. Sin embargo, al evaluar el lexicon en la clase afectiva *Sadness* el rendimiento es malo, y no se producen mejoras al agregar la intensidad afectiva al proceso de análisis.

En definitiva, el lexicon enriquecido tiene un rendimiento promedio de un 52%, con rendimientos aceptables en las clases afectivas *Anger*, *Fear* y *Joy*, y rendimientos deficientes en la clase afectiva *Sadness*.

De los resultados obtenidos, se cree que el rendimiento deficiente en la clase afectiva *Sadness*, tiene directa relación con los bajos índices de correlación que tuvo la misma al evaluar el ordenamiento basado en la intensidad, con las respuestas de los expertos que responden la encuesta y con el conjunto de palabras extraídas de [29]. Se cree que, mejorando los índices de correlación entre la intensidad entregada por el modelo y las respuestas del usuario, es posible mejorar el rendimiento del lexicon.

## Capítulo 7: Trabajo Futuro

En base a todo lo realizado en este trabajo y lo expuesto en el capítulo de discusión, se recomiendan los siguientes trabajos futuros.

- Crear una estrategia que permite definir el extremo de cada clase afectiva y analizar cómo se comporta el modelo propuesto con el extremo afectivo obtenido. ¿Qué pasaría con el modelo de intensidad afectiva, si se obtuviera el extremo afectivo con otra estrategia?
- Modificar el modelo de intensidad afectiva, ya sea utilizando los extremos superior e inferior de cada clase afectiva, o creando otras estrategias para calcular la intensidad a partir de las métricas de similaridad revisadas.
- Generar una estrategia que permita expandir y enriquecer otros lexicones existentes a partir del modelo propuesto en esta investigación utilizando los sinónimos de los términos incluidos en un lexicón, y comprobar si la intensidad afectiva se mantiene entre los sinónimos.
- Desarrollar diferentes modelos de intensidad afectiva, y analizar cómo se comportan en cada clase afectiva, de modo de buscar el mejor modelo para cada clase afectiva.
- Sería interesante desarrollar una aplicación que en base al lexicón enriquecido que se obtiene de esta investigación, pueda analizar textos en internet y que permita obtener retribución de los usuarios, acerca del grado de acierto del análisis realizado.

## Capítulo 8: Conclusiones

En este capítulo se presentan las conclusiones obtenidas a partir de la realización de este trabajo. Primero, se muestran el cumplimiento de la hipótesis y objetivos de este trabajo y luego, las conclusiones generales.

### 8.1. Verificación de Hipótesis y Objetivos

En el capítulo 2 se presentó el objetivo general del trabajo que era: “Enriquecer un lexicón afectivo basado en la clasificación de Plutchik a través del etiquetado automático con la intensidad afectiva de cada palabra utilizando *Wordnet similarity*”. A continuación, se muestran los objetivos específicos y la forma en que se dió cumplimiento a estos.

- Realizar una revisión sistemática de la literatura para seleccionar trabajos anteriores relevantes en el área de análisis de afectos: Se realizó una revisión sistemática de literatura, mediante la cual se pudieron reunir los principales trabajos de creación de lexicones etiquetados con una medida de intensidad afectiva. Además, se analizaron diversos trabajos de creación de lexicones para el análisis de afectos, que comprendieron distintas técnicas y fuentes de información. Esta revisión sistemática permitió contar con un marco teórico de acuerdo al estado del arte en la materia de estudio.
- *Analizar un lexicón de afectos que utilice la clasificación de emociones de Plutchik*: Se revisaron trabajos relacionados a la creación de lexicones de afectos, y se definieron criterios que permitieran seleccionar el lexicón más adecuado según los mismos. De esta selección se obtuvo un lexicón de afectos basado en la clasificación afectiva de Plutchik llamado EmoLex [23] disponible en el sitio web del autor<sup>8</sup>.
- Proponer un modelo que permita incorporar la intensidad afectiva en forma automática, utilizando una métrica de similaridad implementada en *Wordnet Similarity*: Se desarrolló un modelo que permite calcular la intensidad afectiva en forma automática, por medio de las métricas de

---

<sup>8</sup> EmoLex: <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

similaridad implementadas en *Wordnet Similarity*, y se desarrolló una aplicación en el lenguaje Python que a partir de un listado de términos y un extremo de la clase afectiva que se requiere analizar, aplica el modelo de intensidad afectiva para cada término, por medio de las métricas de similaridad revisadas. La aplicación genera 6 salidas en formato .csv, una salida para cada métrica de similaridad. Las salidas se refieren al mismo listado de términos que fueron entregados como entrada, más una nueva columna de datos que se refiere a la medida de intensidad afectiva de cada término.

- *Validar la efectividad de la propuesta a partir de un experimento:* Se diseñaron 2 experimentos, el primero consistió en medir el índice de correlación entre un corpus que se obtuvo desde un lexicón etiquetado con intensidad afectiva descrito en [29], con los resultados de la aplicación del modelo sobre los términos extraídos desde el lexicón [23], enriquecido en esta investigación. Este experimento permitió evaluar el modelo y obtener rendimientos de hasta un 72 % para una de las métricas revisadas. El segundo experimento consistió en evaluar el lexicón de afectos enriquecido mediante el análisis de un corpus de tweets etiquetados con el afecto predominante, los cuales fueron extraídos del trabajo de un investigador en el área del AA. Este experimento entregó un rendimiento general de 52%, y un rendimiento promedio del 62% al realizar el análisis basado en la intensidad afectiva que es añadida al lexicón. Se destaca el rendimiento de 100% para la clase Anger del lexicón al utilizar los valores de intensidad obtenidos con la métrica [16].
- *Analizar objetivamente los resultados obtenidos y generar conclusiones y propuestas de trabajo futuro:* Se presentaron análisis críticos y se compararon los distintos resultados obtenidos del análisis realizado, por clase afectiva y por métrica de similaridad. Estos resultados presentan un aporte para continuar líneas de trabajo futuras relacionadas a la creación y enriquecimiento de lexicones de afectos etiquetados con intensidad afectiva y además para seguir perfeccionando el modelo presentado

De esta forma, mediante estos objetivos específicos se cumplió el objetivo general descrito anteriormente, ya que se enriqueció un lexicón de afectos, añadiendo una medida de intensidad afectiva en forma automática. Esto a través de un modelo que añade la intensidad afectiva automáticamente, por medio de las métricas de similaridad implementadas en *Wordnet Similarity*. Además, se

pudo probar parcialmente la hipótesis descrita en el capítulo 2 “*Un lexicón etiquetado automáticamente con la intensidad afectiva, utilizando la clasificación de emociones de Plutchik, permite mejorar el análisis de afectos basado en lexicón para textos en inglés*” ya que en 3 de las 4 clases evaluadas se obtienen rendimientos considerados aceptables y en 2 de las 3 clases con mejor rendimiento, la medida de intensidad afectiva es fundamental para lograr esos niveles de acierto.

## 8.2. Conclusiones Generales

Mediante este trabajo se enriqueció un lexicón de afectos añadiendo una medida de intensidad afectiva, mediante un modelo que permite incorporar, de forma automática, una medida que indica la intensidad afectiva de una palabra en una clase afectiva, y que utiliza las métricas de similaridad implementadas en *Wordnet*. La revisión sistemática entregó trabajos importantes relacionados con la creación de lexicones de afectos etiquetados con intensidad afectiva. El modelo para incorporación de intensidad afectiva propuesto se basa en el extremo de la clase afectiva y en el uso de las métricas de similaridad implementadas en *Wordnet*. Para probar el modelo se realizó una encuesta que permitió obtener la intensidad, en función del orden de los términos de un subconjunto de 10 palabras por cada clase afectiva que posteriormente se comparó con el ordenamiento entregado por el modelo propuesto. Para el experimento de evaluación del modelo de intensidad afectiva se extrae un corpus de palabras de cada clase afectiva desde un lexicón seleccionado desde los revisados en la revisión sistemática, y se mide el índice de correlación de Pearson de la intensidad, en función del orden de los términos, con el orden entregado por el modelo. En este análisis se obtuvieron resultados de hasta 72% de correlación para la métrica de similaridad Resnik, hasta un 57% para la métrica Jiang&Conrath y un 62 % de correlación general.

Para el experimento de evaluación del lexicón enriquecido se extrae un corpus de tweets etiquetados con el afecto predominando, desde un estudio de un investigador en el área de AA. Se mide el índice *Recall* que indica la cantidad de aciertos respecto al total de documentos evaluados. En este análisis se obtuvieron resultados de hasta un 100% para clase afectiva Anger, utilizando las medidas de

intensidad entregadas por la métrica Lin, con un 62% de rendimiento para el análisis basado en intensidad afectiva, y un 52% de rendimiento general.

A la luz de estos resultados, se concluye que el modelo de intensidad entrega, para algunas clases, resultados aceptables, considerando que es uno de los primeros modelos de cálculo de intensidad afectiva en forma automática. Lo anterior abre bastantes posibilidades de investigación futura, a nivel de mejoras en la forma de decidir el término extremo de la clase afectiva y de la inclusión de nuevas características al modelo de intensidad afectiva. Además, se concluye que el lexicon construido a partir del modelo de intensidad obtiene resultados aceptables para las clases donde los índices de correlación del modelo son aceptables, por lo que las mejoras en el modelo deben impactar positivamente en la calidad de la medida de intensidad del lexicon enriquecido.

## Referencias

1. Pang, B., Lee, L.: Opinion Mining and Sentiment Analysis. *Found. Trends® Inf. Retr.* 2, 1–135 (2008).
2. Plutchik, R.: The emotions. University Press of America (1991).
3. Kensinger, E.A.: Remembering emotional experiences: the contribution of valence and arousal. *Rev. Neurosci.* 15, 241–251 (2004).
4. Ekman, P.: An Argument for Basic Emotions. *Cogn. Emot.* 6, 169–200 (1992).
5. Wilson, T., Wiebe, J., Hoffman, P.: Recognizing contextual polarity in phrase level sentiment analysis. *Acl.* 7, 12–21 (2005).
6. Grefenstette, G., Qu, Y., Shanahan, J., Evans, D.: Coupling Niche Browsers and Affect Analysis for an Opinion Mining Application. *Proc. 12th Int. Conf. Rech. d'Information Assist. par Ordinat.* 186–194 (2004).
7. A Valitutti, C Strapparava, O.S.: Developing Affective Lexical Resources. *PsychNology.* 2, 61–83 (2004).
8. Strapparava, C., Valitutti, A., Stock, O.: The affective weight of lexicon. *Proc. Fifth Int. Conf. Lang. Resour. Eval.* 423–426 (2006).
9. Miller, G.A.: WordNet: a lexical database for English. *Commun. ACM.* 38, 39–41 (1995).
10. Pedersen, T., Patwardhan, S., Michelizzi, J.: WordNet::Similarity - Measuring the relatedness of concepts. In: *Proceedings - Nineteenth National Conference on Artificial Intelligence (AAAI-2004): Sixteenth Innovative Applications of Artificial Intelligence Conference (IAAI-2004).* pp. 1024–1025. Association for Computational Linguistics (2004).
11. Rada, R., Mili, H., Bicknell, E., Blettner, M.: Development and Application of a Metric on Semantic Nets. *IEEE Trans. Syst. Man Cybern.* 19, 17–30 (1989).
12. Wu, Z., Palmer, M.: Verb semantics and lexical selection. *32nd Annu. Meet. Assoc. Comput. Linguist.* 133–138 (1994).
13. Leacock, C., Chodorow, M.: Combining Local Context and WordNet Similarity for Word Sense Identification. *WordNet An Electron. Lex. database.* 265–283 (1998).
14. Resnik, P.: Using Information Content to Evaluate Semantic Similarity in a Taxonomy. *roceedings 14th Int. Jt. Conf. Artif. Intell. - Vol. 1 - IJCAI'95.* 1, 6 (1995).



15. Jiang, J.J., Conrath, D.W.: Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. *Proc. Int. Conf. Res. Comput. Linguist.* 19–33 (1997).
16. Li, Y., Bandar, Z.A., McLean, D.: An approach for measuring semantic similarity between words using multiple information sources. *IEEE Trans. Knowl. Data Eng.* 15, 871–882 (2003).
17. Tversky A.: Features of Similarity. *Psychol. Rev.* 84, 127–190 (1977).
18. Zhou, Z., Wang, Y., Gu, J.: New model of semantic similarity measuring in wordnet. In: *Proceedings of 2008 3rd International Conference on Intelligent System and Knowledge Engineering*. pp. 256–261. IEEE (2008).
19. Meng, L., Huang, R., Gu, J.: A Review of Semantic Similarity Measures in WordNet. *Int. J. Hybrid Inf. Technol.* 6, 1–12 (2013).
20. Ekman, P.: Cross-cultural studies of facial expression. *Darwin facial expressionA century Res. Rev.* 23, 4–5 (1973).
21. Mohammad, S.M., Turney, P.D.: Crowdsourcing the Creation of a Word–Emotion Association Lexicon. *Comput. Intell.* 59, (2008).
22. Mohammad, S.M., Turney, P.D.: Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon. *CAAGET '10 Proc. NAACL HLT 2010 Work. Comput. Approaches to Anal. Gener. Emot. Text.* 26–34 (2010).
23. Mohammad, S.M., Turney, P.D.: Crowdsourcing a word-emotion association lexicon. In: *Computational Intelligence*. pp. 436–465 (2013).
24. Mohammad, S.: #Emotional Tweets. In: *SemEval '12 Proceedings of the First Joint Conference on Lexical and Computational Semantics*. pp. 246–255. Association for Computational Linguistics (2012).
25. Mohammad, S.M., Kiritchenko, S.: Using Hashtags to Capture Fine Emotion Categories from Tweets. *Comput. Intell.* 31, 301–326 (2015).
26. Koto, F., Adriani, M.: HBE: Hashtag-Based Emotion Lexicons for Twitter Sentiment Analysis. In: *Proceedings of the 7th Forum for Information Retrieval Evaluation on - FIRE '15*. pp. 31–34. ACM Press, New York, New York, USA (2015).
27. Volkova, S., Dolan, W.B., Wilson, T., Drive, W.P.: CLex: A Lexicon for Exploring Color , Concept and Emotion Associations in Language. In: *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*. pp. 306–314. ACL (2012).

28. Staiano, J., Guerini, M.: DepecheMood: a Lexicon for Emotion Analysis from Crowd-Annotated News. (2014).
29. Mohammad, S.M.: Word Affect Intensities. (2017).
30. Randolph, J.J.: Free-marginal multirater kappa: An alternative to Fleiss' fixed-marginal multirater kappa. Present. Joensuu Univ. Learn. Instr. Symp. (2005).
31. Fleiss, J.L.: Measuring nominal scale agreement among many raters. Psychol. Bull. 76, 378–382 (1971).
32. Pearson, K.: Note on Regression and Inheritance in the Case of Two Parents. Proc. R. Soc. London. 58, 240–242 (1895).

## ANEXO A. Revisión Sistemática de la literatura

En este anexo se describen los 34 artículos que fueron resultado de la selección de estudios primarios. En la siguiente tabla se describe cada artículo, la fuente, su inclusión/exclusión y comentarios acerca de la exclusión, para los que no fueron seleccionados.

Tabla 18. Estudios incluidos y excluidos de la investigación.

Nº	Estudio	Fuente	Decisión	Comentarios
1	WordNet: a lexical database for English	ACM	Inclusión	Base de datos léxica ampliamente utilizada para el análisis automático de textos
2	Developing Affective Lexical Resources	Psychnology	Inclusión	Expansión de Wordnet que incluye clasificación afectiva. Referenciado en otros artículos sobre análisis afectivo
3	The affective weight of lexicon	Researchgate	Inclusión	Se extiende Wordnet-Affect, incluyendo etiquetas (positiva, negativa, ambigua, neutra)
4	Crowdsourcing the Creation of a Word-Emotion Association Lexicon	Sitio web de Saif Mohammad	Inclusión	Se crea un lexicón para análisis afectivo llamado EmoLex. Se llega a este investigador por referencias de otros trabajos encontrados en ACM
5	Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon	ACM	Inclusión	Mejoras al lexicón EmoLex
6	Crowdsourcing a word-emotion association lexicon	Sitio web de Saif Mohammad	Inclusión	Se mejora el lexicón EmoLex, añadiendo mas términos y nuevas fuentes de información.
7	#Emotional Tweets	ACM	Inclusión	Se construye lexicón específico que relacionan hashtags de twitter con afectos.

Nº	Estudio	Fuente	Decisión	Comentarios
8	Using Hashtags to Capture Fine Emotion Categories from Tweets	ACM	Inclusión	Se construye lexicón específico que relacionan hashtags de twitter con afectos.
9	HBE: Hashtag-Based Emotion Lexicons for Twitter Sentiment Analysis	ACM	Inclusión	Se construye lexicón específico que relacionan hashtags de twitter con afectos.
10	CLex: A Lexicon for Exploring Color	Researchgate	Inclusión	Se construye lexicón que relaciona colores con emociones.
11	DepecheMood: a Lexicon for Emotion Analysis from Crowd-Annotated News	ACM	Inclusión	Se construye un lexicón en forma manual mediante crowdsourcing, incluye la intensidad afectiva de las palabras.
12	Word Affect Intensities	Sitio web de Saif Mohammad	Inclusión	Se construye un lexicón con intensidad afectiva, mediante crowdsourcing. Uso específico para Twitter.
13	Mood patterns and affective lexicon access in weblogs	ACM	Exclusión	No analiza afectos, analiza estados de animo
14	A comprehensive analysis of bilingual lexicon induction	ACM	Exclusión	Traduce un lexicón, no crea ni hace un análisis basado en lexicón
15	Build Emotion Lexicon from Microblogs by Combining Effects of Seed Words and Emoticons in a Heterogeneous Graph	ACM	Exclusión	Propuesta de framework para construir lexicones
16	Seeing the best and worst of everything on the web with a two-level, feature-rich affect lexicon	ACM	Exclusión	Propuesta para modificar el afecto de un concepto según contexto
17	Research on web topic detection based on domain lexicon	IEEE	Exclusión	Propone framework para crear lexicón de dominio específico
18	Lexicon Generation for Emotion Detection from Text	IEEE	Exclusión	Modelo para actualizar lexicones y adaptarlos a dominios específicos
19	A Hybrid Sentiment	IEEE	Exclusión	Análisis de sentimientos, no de

Nº	Estudio	Fuente	Decisión	Comentarios
	Lexicon for Social Media Mining			afectos
20	Construction and quantization for a basic sentiment lexicon	IEEE	Exclusión	Creación de lexicones de sentimientos, no de afectos
21	Real time analysis of top trending event on Twitter: Lexicon based approach	IEEE	Exclusión	Análisis de sentimientos, no de afectos
22	Sentiment analysis on facebook group using lexicon based approach	IEEE	Exclusión	Análisis de sentimientos, no de afectos
23	Lexicon based feature extraction for emotion text classification	ScienceDirect	Exclusión	Modelo para actualizar lexicones
24	Web Questionnaire as Construction Method of Affect-annotated Lexicon - Risks Reduction Strategy	IEEE	Exclusión	No construye lexicón, propone técnicas para disminuir los riesgos al crear lexicones basados en encuestas.
25	Affect Analysis of Web Forums and Blogs Using Correlation Ensembles	IEEE	Exclusión	Comparativa del rendimiento de distintos lexicones de afectos.
26	Building a Twitter opinion lexicon from automatically-annotated tweets	ScienceDirect	Exclusión	Construye un lexicón de sentimientos, no de afectos.
27	Improvement of an emotional lexicon for the evaluation of beers	ScienceDirect	Exclusión	El lexicón generado tiene sólo 28 palabras y no se basa en clasificación afectiva conocida. Su experimento no entrega valor a este estudio
28	From emotion to language: Application of a systematic, linguistic-based approach to design a food-associated emotion lexicon	ScienceDirect	Exclusión	Construye un lexicón de sentimientos, no de afectos.
29	Sentiment and emotion classification over noisy	ScienceDirect	Exclusión	Se propone técnica para mejorar la calidad de la clasificación de

Nº	Estudio	Fuente	Decisión	Comentarios
	labels			los términos de los lexicones.
30	Building Emotion Lexicon from Weblog Corpora	ACM	Exclusión	Construye un lexicón de sentimientos, no de afectos.
31	Using Google n-Grams to Expand Word-Emotion Association Lexicon	Springer	Exclusión	Propone estrategia para mejorar un lexicón de un estudio ya incluido en este informe
32	Emotion classification using massive examples extracted from the web	ACM	Exclusión	La propuesta es construida para idioma japonés.
33	Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text	ScienceDirect	Exclusión	Revisión Sistemática de la literatura sobre análisis de afectos y de sentimientos.
34	A data-driven classification of feelings	ScienceDirect	Exclusión	Construye un lexicón de sentimientos, no de afectos.

## ANEXO B. Encuesta de opinión sobre intensidad afectiva

Este anexo describe el entorno de desarrollo y los métodos utilizados para implementar la encuesta de opinión sobre intensidad afectiva. Esto, mediante una aplicación web.

### B.1. Entorno de desarrollo

#### Características del Hardware

Durante el desarrollo de la aplicación web se utilizó un computador Apple MacBook Air con las siguientes características:

- Procesador Intel Core i5 2.7GHz
- 8 GB de memoria RAM
- 256GB SSD
- Sistema Operativo MacOS Sierra 10.12

#### Plataforma de desarrollo

La aplicación web desarrollada, utiliza como plataforma principal PHP para el funcionamiento del servidor, y HTML, CSS y JavaScript para el funcionamiento en el navegador web del usuario, y PostgreSQL como motor de bases de datos. A continuación, se describen las librerías utilizadas en su desarrollo:

- PostgreSQL<sup>9</sup>: Motor de bases de datos relacionales, utilizado para mantener la información de la encuesta.
- PHP<sup>10</sup>: Lenguaje de programación interpretado, utilizado principalmente para servidores web. En esta aplicación se utiliza principalmente como un controlador de solicitudes que vienen desde los navegadores web de los

---

<sup>9</sup> <https://www.postgresql.org/>

<sup>10</sup> <http://php.net/>

usuarios encuestados, y recuperando o almacenando información en la base de datos.

- HTML: Lenguaje de etiquetas para dar estructura a un sitio web
- CSS: Lenguaje usado para definir la presentación de una página web. En esta aplicación se utiliza para manejar la compatibilidad con dispositivos móviles. Esto, para permitir que los usuarios tengan más posibilidades de contestar la encuesta.
- Javascript: Lenguaje interpretado, utilizado tanto en clientes como servidores. En esta aplicación se utiliza para controlar las validaciones necesarias antes de recuperar o enviar datos al servidor web.
- JQuery<sup>11</sup>: Biblioteca que facilita la interacción con eventos de una aplicación web y permite añadir mejoras visuales. En esta aplicación se utiliza para simplificar la forma de realizar peticiones de información hacia el servidor web.
- JQuery UI<sup>12</sup>: Biblioteca de plugins extras para JQuery. En esta aplicación se utiliza para controlar el efecto que permite ordenar las palabras arrastrando el mouse o usando el dedo en dispositivos táctiles.

## Base de Datos

Para almacenar la información referente a las palabras de cada clase afectiva y las respuestas de los usuarios que responden la encuesta se utiliza una base de datos PostgreSQL con la siguiente estructura.

---

<sup>11</sup> <https://jquery.com/>

<sup>12</sup> <https://jqueryui.com/>



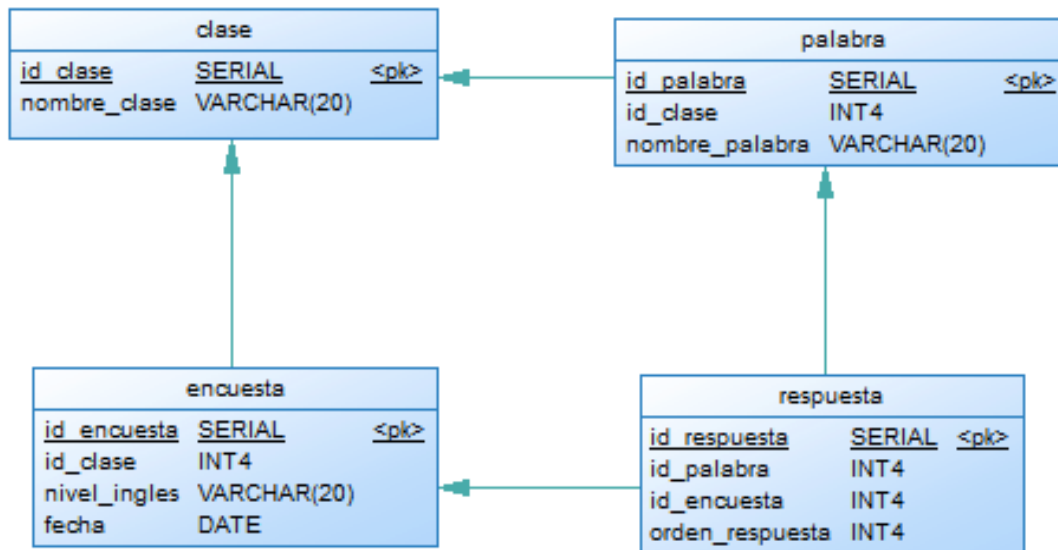


Figura 8-1. Base de datos para la encuesta de opinión.

La Base de datos consiste de las siguientes 4 tablas:

- Clase: Se utiliza para almacenar cada una de las 8 clases afectiva de Plutchik.
- Palabra: Se utiliza para almacenar las palabras seleccionadas para ser incluidas en la encuesta (10 palabras por cada clase afectiva). Se mantiene una referencia de la clase a la que pertenece la palabra.
- Encuesta: Se utiliza para almacenar información de la encuesta que responden los usuarios, específicamente, el nivel de inglés del usuario y la fecha en que responde. Además se mantiene una referencia a la clase para la que realiza sus respuestas el usuario.
- Respuesta: Se utiliza para almacenar las respuesta de cada encuesta contestada por los usuarios. Se hace referencia a la palabra que contestó el usuario, y el orden de intensidad en que la catalogó.

## B.2. Funcionamiento de la aplicación encuesta de opinión

A continuación se describen los pasos que permiten contestar una encuesta de opinión acerca de la intensidad afectiva.

La encuesta de opinión se encuentra disponible en el sitio web<sup>13</sup> de grupo de investigación SOMOS de la Universidad del Bío-Bío. Al acceder se visualiza la aplicación.

- 1 – Selección de Clase Afectiva: permite seleccionar la clase afectiva para la que se desea responder la encuesta
- 2 – Nivel de inglés: permite al encuestado, seleccionar el nivel de inglés que posee. Esta información se considera relevante para identificar las respuestas de los usuarios, y de ser posible, utilizar solo las respuestas de usuarios con nivel avanzado o nativo.
- 3 – Responder Encuesta: Al seleccionar la Clase afectiva, esta sección se llena con 10 palabras que aparecen en el rectángulo blanco, las que deben ser arrastradas al rectángulo amarillo según la intensidad del afecto, según la opinión del usuario.
- 4 – Enviar Encuesta: Una vez que la encuesta fue contestada puede ser enviada. Si falta información por completar, se cancela el envío y se informa al usuario para que corrija antes de volver a enviar.

### B.3. Resultados de la encuesta de opinión

Una vez que se cierra el período habilitado para responder la encuesta, se construyen consultas que permitan extraer las respuestas desde la base de datos. El formato de las respuestas que se obtienen es el siguiente, obtenido de una encuesta real:

Encuesta n° 1: Clase Anger. Nivel de inglés avanzado

Tabla 19. Formato de respuesta de encuesta real

Palabra	Orden
discontent	0
upset	1
indignation	2
frustration	3
irritation	4

<sup>13</sup> <http://somos.ubiobio.cl/encuesta/>

Palabra	Orden
resentment	5
wrath	6
hostility	7
fury	8
rage	9

Con este mismo formato de resultados, se obtienen las respuestas de todas las encuestas, para la clase Anger se obtuvieron 6 respuestas, para la clase Anticipation se obtuvieron 3 respuestas, para la clase Disgust se obtuvieron 4 respuestas, para la clase Fear se obtuvieron 3 respuestas, para la clase Joy se obtuvieron 8 respuestas, para la clase Sadness se obtuvieron 5 respuestas, para la clase Trust se obtuvieron 3 respuestas, y para la clase Surprise se obtuvieron 3 respuestas. El detalle de respuestas por clase y por nivel de inglés se presenta en la siguiente Tabla:

Tabla 20. Encuestas por clase afectiva y nivel de inglés

Clase	Nivel básico	Nivel intermedio	Nivel avanzado	Nivel Nativo
Anger	0	0	5	1
Anticipation	0	2	0	1
Disgust	0	1	2	1
Fear	0	1	1	1
Joy	0	1	6	1
Sadness	0	1	3	1
Trust	0	1	2	0
Surprise	0	1	2	0

## ANEXO C. Implementación

### C.1. Aplicación para el cálculo de la intensidad afectiva

Este anexo describe el entorno de desarrollo y los métodos utilizados para implementar la etapa de cálculo del modelo de intensidad afectiva propuesto en el trabajo, mediante una aplicación software. Esta aplicación permite realizar tareas de cálculo de la intensidad afectiva de un conjunto de términos y construcción de un listado de palabras etiquetado con intensidad afectiva, en formato .csv.

#### Características del hardware

Durante el desarrollo de la aplicación, se utilizó un computador Apple MacBook Air con las siguientes características:

- Procesador Intel Core i5 2.7GHz
- 8 GB de memoria RAM
- 256GB SSD
- Sistema Operativo MacOS Sierra 10.12

#### Entorno de desarrollo

La aplicación desarrollada utiliza como plataforma principal Python en su versión 2.7.10 y un conjunto de librerías de procesamiento de texto los cuales se detallan a continuación:

- NLTK<sup>14</sup>: Natural Language Toolkit, es una librería de Python que permite realizar análisis y procesamiento lingüístico. En este trabajo utilizamos NLTK principalmente para aprovechar su implementación de las métricas de similaridad implementadas en Wordnet Similarity.
- NLTK.wordnet: es un sub-módulo de NLTK que contiene una implementación de Wordnet para Python.

---

<sup>14</sup> <https://www.nltk.org/>

- NLTK.wordnet\_ic: es un sub-módulo de NLTK que contiene una implementación que permite obtener Corpus de Wordnet con Contenido Informativo (IC), necesarios para realizar el cálculo de similaridad con algunas métricas de Wordnet Similarity.
- CSV: librería incluida en Python, que provee la funcionalidad necesaria para generar archivos .csv a partir de arrays de Python.

## Implementación del cálculo de intensidad afectiva

Se describen a continuación las funciones utilizadas para realizar el cálculo de intensidad afectiva:

- Configuración inicial: En esta etapa se carga un corpus de Wordnet, utilizando el método **ic()** provisto por NLTK.wordnet\_ic. Se obtiene el synset de el término que representa al extremo superior (mayor intensidad) de la clase afectiva mediante el método **synset()** provisto por NLTK.wordnet. Por último se construyen los arrays con el listado de términos a etiquetar y para cada una de las métricas de similaridad a utilizar en el cálculo.
- Cálculo de similaridad: En esta etapa se obtiene el synset de cada término a etiquetar, utilizando el método **synset()** provisto por NLTK.wordnet. Luego, se realiza el cálculo de similaridad entre el synset de el extremo superior de la clase afectiva y cada término a etiquetar. Esto, mediante los métodos **wup\_similarity()**, **path\_similarity()**, **lch\_similarity()**, **res\_similarity()**, **jcn\_similarity()** y **lin\_similarity()**. Todos ellos provistos por NLTK.wordnet. Cada uno de estos cálculos se almacena en un array distinto, que representa a cada una de las métricas de similaridad.
- Exportación de resultados: En esta etapa se utiliza la librería CSV provista por Python, donde se construye un archivo CSV para cada array, y se exportan utilizando el método **writer()** de CSV.

## Resultados al ejecutar la aplicación

Los resultados de la aplicación, se presentan en formato csv, y tienen un formato que consiste en la palabra, y su valor de similaridad con el extremo

superior de la clase afectiva. En la siguiente tabla se presenta un extracto de la información.

Clase afectiva: Anger

Extremo superior: rage

Métrica: wup

Tabla 21. Ejemplo de resultados del cálculo de intensidad

Palabra	Similaridad con extremo
fury	1,0
wrath	0,9412
anger	0,9333
exasperation	0,8236

## C.2. Aplicación para experimento de evaluación

Este anexo describe el entorno de desarrollo y los métodos utilizados para implementar el experimento de evaluación. Esto, mediante una aplicación web.

### Entorno de desarrollo

#### Características del Hardware

Durante el desarrollo de la aplicación web se utilizó un computador Apple MacBook Air con las siguientes características:

- Procesador Intel Core i5 2.7GHz
- 8 GB de memoria RAM
- 256GB SSD
- Sistema Operativo MacOS Sierra 10.12

#### Plataforma de desarrollo

La aplicación web desarrollada, funciona totalmente en el navegador del usuario, por lo que utiliza como plataforma principal Javascript para las operaciones de carga de archivos y cálculo de resultados. Además se utiliza HTML y CSS. A continuación se describen las librerías utilizadas en su desarrollo:

- HTML: Lenguaje de etiquetas para dar estructura a la aplicación web
- CSS: Lenguaje usado para definir la presentación de una página web. En esta aplicación se utiliza para manejar la compatibilidad con distintos

tamaños de pantalla. Esto, para permitir que los usuarios tengan más posibilidades de contestar la encuesta.

- **Javascript:** Lenguaje interpretado, utilizado tanto en clientes como servidores. En esta aplicación se utiliza para cargar los archivos necesarios, y para realizar la ejecución del proceso de validación.
- **jQuery<sup>15</sup>:** Biblioteca que facilita la interacción con eventos de una aplicación web y permite añadir mejoras visuales. En esta aplicación se utiliza para simplificar la forma de cargar los resultados del experimento en el navegador.

## Funcionamiento de la aplicación experimento de evaluación

A continuación se describe la forma de utilizar la aplicación para el experimento de evaluación.

La encuesta de opinión se encuentra disponible dentro del sitio web<sup>16</sup> del grupo de investigación SOMOS de la Universidad del Bío-Bío. Al acceder se visualiza la siguiente aplicación.

**Análisis Afectivo** IntensidadAutomática

Esta aplicación permite analizar afectivamente un listado de textos

Configuración **1** Resultados

---

**Configuración** **2**

Antes de iniciar el análisis debe subir los archivos que se solicitan, incluyendo lexicón etiquetado con intensidad afectiva y textos a analizar

Formato lexicón: [palabra].[tipo].[pos]:[intensidad]\n

Formato textos: [cod]t[frase]t[emoción]t[intensidad]

Métrica	Lexicón Anger <b>3</b>	Lexicón Fear	Lexicón Sadness	Lexicón Joy
PATH	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo
WUP	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo
LCH	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo
RES	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo
JCN	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo
LIN	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo	<input type="button" value="Seleccionar archivo"/> No se eligió archivo

Texto a analizar **4**

Stopwords  No se eligió archivo

Cobertura Mínima  No se eligió archivo

0 %

**Comenzar Análisis** **5**

Figura 8-2. Aplicación web, screenshot experimento de evaluación

<sup>15</sup> <https://jquery.com/>

<sup>16</sup> <http://somos.ubiobio.cl/analisis/>

De la figura se distinguen algunos elementos principales, destacados con números rojos:

- 1 – Pestañas de opción: Permite seleccionar el contenido que se desea visualizar, en un inicio se ve el contenido de la pestaña Configuración, la que una vez utilizada permite cambiar a la pestaña Resultados, para ver el análisis en detalle.
- 2 – Instrucciones: Entrega información relevante acerca de los formatos necesarios para el correcto funcionamiento de la aplicación.
- 3 – Carga de archivos: Permite cargar los archivos con los lexicones a utilizar. A modo de hacer un análisis más sencillo, se separa la carga por clase afectiva y por métrica utilizada. Es importante destacar que solo es necesario subir los archivos que queremos utilizar para el análisis. Los archivos de clases y métricas no que no se suben a la aplicación, indican que no requerimos realizar un análisis con esos criterios.
- Configuración: Consta de 3 etapas requeridas para el análisis
  - Texto a analizar: Permite subir el archivo con el listado de textos que deseamos analizar utilizando nuestro lexicón.
  - Stopwords: Permite subir un archivo con un listado de palabras que serán ignoradas en el análisis afectivo. Para el análisis descrito en el informe se utilizó un listado <sup>17</sup> de stopwords creado por la universidad de Standford, y representa un estándar para el análisis de textos.
  - Porcentaje de cobertura: Permite indicar a la aplicación, el porcentaje de palabras que deben ser analizadas para considerar un análisis como válido. Para el experimento desarrollado en este informe, se indica un 33 por ciento de cobertura mínima.
- Análisis: Permite comenzar el análisis del archivo seleccionado, utilizando los lexicones cargados, considerando los stopwords y la cobertura mínima indicada.

---

17

<https://github.com/stanfordnlp/CoreNLP/blob/master/data/edu/stanford/nlp/patterns/surface/stopwords.txt>



## Implementación de la aplicación de análisis afectivo

Se describen a continuación las funciones utilizadas y construidas para realizar el análisis afectivo:

- FileReader: Implementación de lectura de archivos asíncrona integrada en Javascript. En esta aplicación se utiliza para cargar los archivos de lexicón, el texto a analizar y los stopwords. Una vez cargados estos archivos, se almacena cada palabra etiquetada en estructuras **Map(clave, valor)**, utilizando el método **readAsText()**.
- Análisis: Esta etapa es construida utilizando Javascript, sin librerías o módulos adicionales. Se divide en 4 etapas que se describen a continuación:
  - Tokenización: Esta etapa consiste en descomponer un elemento del corpus (frase) en distintas partes, utilizando el método **split()** provisto por Javascript. Esta función permite mantener en un arreglo cada parte de la frase para un posterior análisis.
  - Filtrado de caracteres: Mediante la función **replace()** es posible filtrar cada token que no sea una palabra (emoticonos, hipervínculos, hashtags, etc) y entregar una versión sanitizada de cada elemento a analizar.
  - Verificación: En esta etapa se verifica que cada token sea candidato a ser analizado. Esto, cuando no se encuentra en el listado de stopwords definidos.
  - Análisis: En esta etapa se analiza cada una de las palabras verificadas y sanitizadas. Se almacena el resultado de su búsqueda dentro del lexicón para el análisis por densidad de palabras y por intensidad de palabras.
  - Cálculo: En esta etapa se realizan los cálculos que permiten entregar los resultados para el análisis por densidad e intensidad de palabras.
- Resultados: En esta etapa se utiliza el método **each()** de jQuery para presentar los resultados de forma visual en la pestaña correspondiente. Una vez finalizada esta etapa se le informa al usuario que el análisis ha finalizado y que puede ver los resultados.

## Presentación de resultados

Una vez completada la etapa de análisis, se presentan los resultados en una tabla que visualiza continuación.

**Análisis Afectivo** IntensidadAutomática

Esta aplicación permite analizar afectivamente un listado de textos

Configuración Resultados

**Resultados del Análisis**

Legenda: Anger Fear Sadness Joy

Nº	Frase	Afecto - Intensidad	PATH	WUP	LCH	RES	JCN	LIN
0	I feel like I am drowning. #depression #anxiety #failure #worthless	fear - 0.979	Afecto: Sadness 0.980 0.990 0.990 Cobertura: 40.0%	Afecto: Sadness 0.980 0.980 0.780 Cobertura: 40.0%	Afecto: Sadness 0.981 0.981 0.940 Cobertura: 40.0%	Afecto: Sadness 0.980 0.980 0.980 Cobertura: 40.0%	Afecto: Sadness 0.980 0.980 0.974 Cobertura: 40.0%	Afecto: Sadness 0.981 0.981 0.980 Cobertura: 40.0%
2	I lost my blinders .... #panic	fear - 0.975	Afecto: Fear 0.990 0.990 0.990 Cobertura: 16.7%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 16.7%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 16.7%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 16.7%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 16.7%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 16.7%
3	I feel like I am drowning. #depression #failure #worthless	fear - 0.938	Afecto: Sadness 0.990 0.990 0.990 Cobertura: 10.0%	Afecto: Sadness 0.990 0.990 0.990 Cobertura: 10.0%	Afecto: Sadness 0.990 0.990 0.990 Cobertura: 10.0%	Afecto: Sadness 0.990 0.990 0.990 Cobertura: 10.0%	Afecto: Sadness 0.990 0.990 0.990 Cobertura: 10.0%	Afecto: Sadness 0.990 0.990 0.990 Cobertura: 10.0%
4	This is the scariest American Horror Story out of all of them... I'm gonna have to watch in the daytime. #frightened	fear - 0.938	Afecto: Fear 0.990 0.990 0.990 Cobertura: 14.3%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 14.3%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 14.3%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 14.3%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 14.3%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 14.3%
5	@mgcasarwork I nearly started crying and having a full on panic attack after tatino bc of the crowds so I feel him	fear - 0.938	Afecto: Fear 0.990 0.990 0.990 Cobertura: 4.5%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 4.5%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 4.5%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 4.5%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 4.5%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 4.5%
6	I have to finally tell my therapist about my sexuality ... last frontier ... not sure I can do it in the AM #fear #SingleGirlProblems	fear - 0.938	Afecto: Fear 0.990 0.990 0.990 Cobertura: 8.0%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 8.0%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 8.0%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 8.0%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 8.0%	Afecto: Fear 0.990 0.990 0.990 Cobertura: 8.0%

Figura 8-3 afectivo. Tabla de resultados, screenshot aplicación web de análisis afectivo

De esta tabla se destacan los siguientes elementos:

- **Leyenda:** En esta sección se presentan 4 colores que representan a las clases afectivas analizadas: Anger (rojo), Fear (naranja), Sadness (celeste) y Joy (verde).
- **Frase:** Esta columna muestra la frase original que fue analizada, se incluyen caracteres sin limpiar y stopwords.
- **Afecto – Intensidad:** Esta columna indica el afecto y la intensidad originales que tiene la frase, según el trabajo [29]
- **Resultados (PATH, WUP, LCH, RES, JCN, LIN):** Cada una de estas 6 columnas representa los resultados de una métrica. Se indica el afecto predominando según la estrategia densidad de palabras, y se presentan 4 leyendas que contienen los valores de intensidad para cada una de las clases analizadas. Estas leyendas representan a el análisis por la estrategia intensidad de palabras. Además, se incluye información acerca del porcentaje de cobertura del análisis realizado.

## ANEXO D. Anexo Matriz y análisis Kappa.

Tabla 22. Matriz de respuestas a la encuesta para la Clase *Anger*

Palabra / Clasificación	1	2	3	4	5
discontent	6	0	0	0	0
upset	5	1	0	0	0
resentment	0	2	2	2	0
irritation	0	2	3	1	0
indignation	0	2	3	1	0
frustration	1	4	1	0	0
hostility	0	0	2	3	1
wrath	0	1	1	1	3
fury	0	0	0	1	5
rage	0	0	0	3	3
Overall Agreement	43%				
Fixed Marginal Kappa	0,29				
Free Marginal Kappa	0,29				

Tabla 23. Matriz de respuestas a la encuesta para la Clase *Anticipation*

Palabra / Clasificación	1	2	3	4	5
fancy	2	0	1	0	0
whim	2	0	1	0	0
aspiration	0	1	0	2	0
craving	1	1	0	0	1
expectation	0	2	0	1	0
eagerness	0	0	1	1	1
excitement	0	0	1	0	2
ambition	0	0	0	1	2
optimism	1	1	0	1	0
hopefulness	0	1	2	0	0
Overall Agreement	23%				
Fixed Marginal Kappa	0,04				
Free Marginal Kappa	0,04				

Tabla 24. Matriz de respuestas a la encuesta para la Clase *Disgust*

Palabra / Clasificación	1	2	3	4	5
dislike	1	2	1	0	0
distaste	2	2	0	0	0
disgust	1	0	3	0	0
displeasure	2	2	0	0	0
aversion	1	2	1	0	0
revulsion	0	0	0	3	1
hatred	0	0	0	2	2
horror	0	0	0	2	2
animosity	1	0	3	0	0
loathing	0	0	0	1	3
Overall Agreement	37%				
Fixed Marginal Kappa	0,21				
Free Marginal Kappa	0,21				

Tabla 25. Matriz de respuestas a la encuesta para la Clase *Fear*

Palabra / Clasificación	1	2	3	4	5
worry	3	0	0	0	0
alarm	3	0	0	0	0
scare	0	0	1	1	1
anxiety	0	3	0	0	0
trepidation	0	1	1	0	1
fright	0	2	0	1	0
fear	0	0	3	0	0
dread	0	0	1	2	0
panic	0	0	0	2	1
terror	0	0	0	0	3
Overall Agreement	60%				
Fixed Marginal Kappa	0,50				
Free Marginal Kappa	0,50				

Tabla 26. Matriz de respuestas a la encuesta para la Clase *Joy*

Palabra / Clasificación	1	2	3	4	5
amusement	5	1	0	1	1
gladness	5	3	0	0	0
delight	1	3	3	1	0
cheerfulness	3	3	2	0	0
happiness	0	3	3	2	0
joy	0	1	5	2	0
rejoicing	1	0	1	6	0
glee	1	2	2	1	2
exhilaration	0	0	0	3	5
ecstasy	0	0	0	0	8
Overall Agreement	40%				
Fixed Marginal Kappa	0,25				
Free Marginal Kappa	0,25				

Tabla 27. Matriz de respuestas a la encuesta para la Clase *Sadness*

Palabra / Clasificación	1	2	3	4	5
unhappiness	1	3	0	1	0
pity	3	1	1	0	0
distress	1	0	2	2	0
regret	2	1	2	0	0
melancholy	2	0	2	1	0
sorrow	1	2	2	0	0
depression	0	0	0	2	3
suffering	0	0	0	3	2
agony	0	0	0	0	5
grief	0	2	2	1	0
Overall Agreement	34%				
Fixed Marginal Kappa	0,17				
Free Marginal Kappa	0,18				

Tabla 28. Matriz de respuestas a la encuesta para la Clase *Trust*

Palabra / Clasificación	1	2	3	4	5
liking	2	1	0	0	0
leaning	1	1	0	1	0
respect	0	0	3	0	0
approvation	0	2	0	1	0
optimism	2	0	1	0	0
confidence	0	0	0	1	2
loyalty	0	0	0	1	2
reverance	0	0	0	1	2
hope	1	2	0	0	0
admiration	0	0	2	1	0
Overall Agreement	37%				
Fixed Marginal Kappa	0,21				
Free Marginal Kappa	0,21				

Tabla 29. Matriz de respuestas a la encuesta para la Clase *Surprise*

Palabra / Clasificación	1	2	3	4	5
stab	1	0	0	2	0
startle	1	1	1	0	0
jolt	1	2	0	0	0
scare	0	1	1	1	0
surprise	2	0	1	0	0
excitement	0	2	1	0	0
astonishment	0	0	1	1	1
amaze	1	0	0	2	0
stunned	0	0	0	0	3
shock	0	0	1	0	2
Overall Agreement	30%				
Fixed Marginal Kappa	0,13				
Free Marginal Kappa	0,13				