Escuela Politécnica Superior de Jaén

Grado en Ingeniería Informática

**UNIVERSIDAD DE JAÉN**
*Escuela Politécnica Superior de Jaén*

Trabajo Fin de Grado

# GENERACIÓN DE UN RECURSO LÉXICO PARA EL ANÁLISIS DE EMOCIONES EN ESPAÑOL

**Alumna: María del Mar Rojas Moya**

Tutoras: Prof. Dña. Salud Mª Jiménez Zafra
Prof. Dña. Mª Dolores Molina González
Dpto: Informática

**Octubre, 2018**

Universidad de Jaén
Escuela Politécnica Superior de Jaén
Departamento de Informática

Doña Salud Mª Jiménez Zafra y Doña Mª Dolores Molina González, tutoras del Proyecto Fin de Grado titulado: Generación de un recurso léxico para el análisis de emociones en español, que presenta María del Mar Rojas Moya, autorizan su presentación para defensa y evaluación en la Escuela Politécnica Superior de Jaén.

Jaén, OCTUBRE de 2018

La alumna:

ROJAS MOYA MARÍA DEL MAR - 77384203J

Firmado digitalmente por ROJAS MOYA MARÍA DEL MAR - 77384203J
Fecha: 2018.10.30 09:52:24 +01'00'

Fdo: María del Mar Rojas Moya

Las tutoras:

JIMENEZ ZAFRA SALUD MARIA - 77370897R

Firmado digitalmente por JIMENEZ ZAFRA SALUD MARIA - 77370897R
Nombre de reconocimiento (DN): c=ES, serialNumber=IDCES-77370897R, givenName=SALUD MARIA, sn=JIMENEZ ZAFRA, cn=JIMENEZ ZAFRA SALUD MARIA - 77370897R
Fecha: 2018.10.30 09:23:41 +01'00'

Fdo: Prof. Dña. Salud Mª Jiménez Zafra

MOLINA GONZALEZ MARIA DOLORES - 26009191D

Firmado digitalmente por MOLINA GONZALEZ MARIA DOLORES - 26009191D
Nombre de reconocimiento (DN): c=ES, serialNumber=26009191D, sn=MOLINA GONZALEZ, givenName=MARIA DOLORES, cn=MOLINA GONZALEZ MARIA DOLORES - 26009191D
Fecha: 2018.10.30 09:33:24 +01'00'

Fdo: Prof. Dña. Mª Dolores Molina González

# Índice

## Resumen

El Trabajo de Fin de Grado que se va a exponer a continuación lidia con el objetivo de generar un recurso léxico en español que pueda ser utilizado para clasificación de emociones. A consecuencia de ello, también se creará una herramienta para poder analizar emociones y diversas pruebas realizadas sobre corpus etiquetados con esta herramienta para poder comprobar el rendimiento de la misma. A lo largo de cada uno de estos pasos, se hablará sobre algunos conceptos teóricos fundamentales para entender la base de este proyecto y así entender los diferentes modos de proceder que se han seguido a la hora de la implementación de cada una de las partes de proyecto.

# 1. INTRODUCCIÓN.

## 1.1. Motivación.

En la actualidad, con la revolución que han supuesto las nuevas tecnologías, podemos ver cómo se pueden utilizar todos estos elementos que han surgido para hacernos la vida más fácil y ayudarnos en diferentes y muy dispares situaciones.

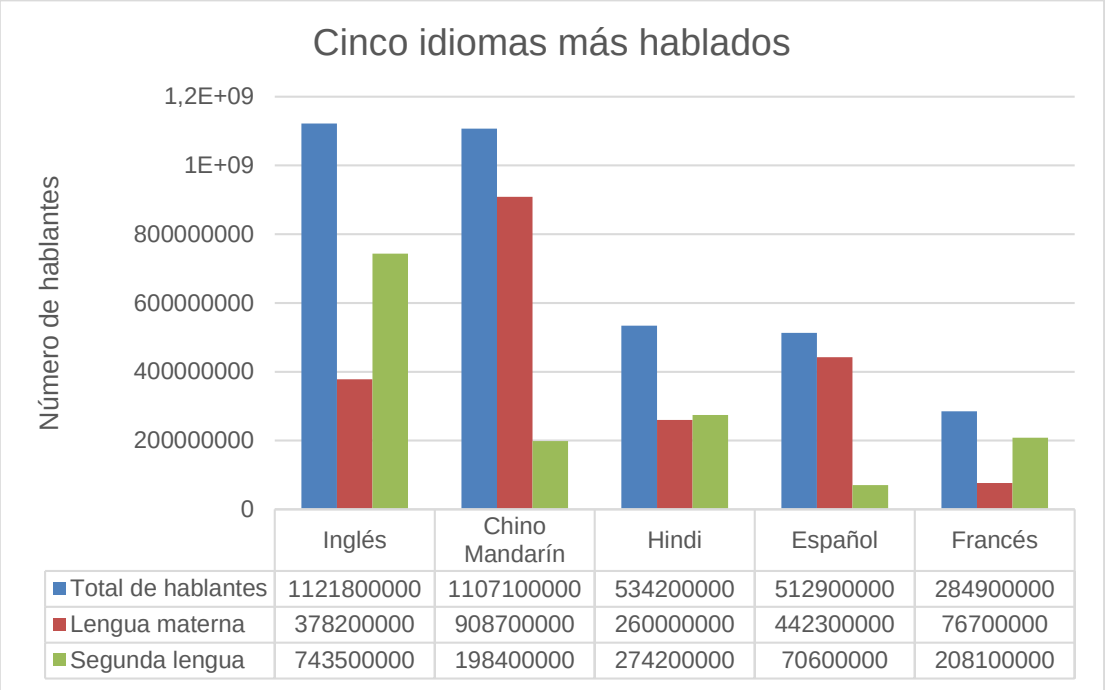
Nos podemos encontrar con una amplia variedad de campos de estudio cuya aplicabilidad es altamente interesante en la sociedad. Un elemento de gran atracción y presumiblemente novel de esta revolución, y en el que se va a centrar el presente trabajo, es el análisis de emociones, siendo un procedimiento que tiene múltiples aplicaciones y que puede resultar de gran utilidad en determinadas situaciones clave.

El rango de operación de este campo de estudio es muy amplio, y podemos encontrar aplicaciones tanto para la personalización en sistemas de recomendación basados en emociones, como podría ser la música, cierto tipo de respuestas por parte de agentes de voz con los que a día de hoy tratamos de manera diaria, como el Asistente de Google, Siri, Cortana o Alexa, u otros casos de mayor relevancia, como la detección de tendencias suicidas en una persona a través de la detección de las emociones en su discurso.

Como la utilidad de esta herramienta es bastante alta y puede resultar harto ventajosa, y la existencia de emociones en nuestro discurso es algo que se produce de manera bastante frecuente, surge la idea de crear un léxico que nos permita tratar estas emociones y poder estudiarlas correctamente, para después proceder a realizar, en base a los resultados obtenidos, las acciones pertinentes.

Si bien actualmente contamos con una amplia variedad de herramientas que nos permiten realizar este estudio, la gran mayoría de ellas no se encuentran en idioma español, si no que se encuentran en inglés, lo que supone una barrera para el correcto desarrollo de estos nuevos procedimientos de tratamiento de emociones en nuestra lengua. Por tanto, surge la necesidad de tener estos mismos recursos pero para el idioma español, y así poder abrir este campo de análisis a toda la comunidad

hispanoparlante, que, como podemos apreciar en la ilustración 1, es una porción muy significativa [1]:




| | Inglés | Chino Mandarín | Hindi | Español | Francés |
|---|---|---|---|---|---|
| Total de hablantes | 1121800000 | 1107100000 | 534200000 | 512900000 | 284900000 |
| Lengua materna | 378200000 | 908700000 | 260000000 | 442300000 | 76700000 |
| Segunda lengua | 743500000 | 198400000 | 274200000 | 70600000 | 208100000 |

**Ilustración 1. Gráfica de los cinco idiomas más hablados del mundo**

La información que arrojan los datos es clara: el español es la cuarta lengua más hablada teniendo en cuenta el total de hablantes, y es la segunda como lengua materna, por dentrás del chino mandarín. A la luz de los resultados obtenidos, surge la necesidad de la generación de un recurso léxico que pueda ser de utilidad para este conjunto de personas.

## 1.2. Propósito

Para poder lograr este análisis de emociones en español, se va a proceder a la utilización de diferentes recursos léxicos en inglés para su correspondiente traducción y uso en español. De hecho, se irá un paso más allá para proceder a un estudio de diversas fuentes a través de variadas herramientas y técnicas, para poder asegurar la diversidad de los resultados.

Para ello, y para la correcta elección de los mismos, se realizará el consecuente estudio de las emociones y de los modelos emocionales que existen en la actualidad. Veremos también los distintos medios que podemos utilizar y el por qué de utilizarlos, además de explicar de una manera clara y detallada qué procedimientos se han llevado a cabo para la obtención de la traducción de los recursos. Por último, para poner a prueba lo creado, se realizará un pequeño programa de análisis de emociones que nos permitirá ver en funcionamiento la herramienta generada.

## 1.3. Objetivos

Para el correcto desempeño del proyecto, a continuación se definen los objetivos o requisitos que resultan fundamentales para el proyecto y cuya consecución conlleva la realización correcta y completa del presente proyecto:

1. Realizar un estudio de los lexicones etiquetados con emociones que existen en español y en inglés.
2. Estudiar diferentes aproximaciones para la generación de recursos para el análisis de emociones.
3. Generar un recurso léxico para el análisis de emociones en español.
4. Estudiar diferentes técnicas para la clasificación de emociones en textos.
5. Implementar una herramienta que incorpore alguno de los métodos de clasificación estudiados para mostrar la utilidad del recurso generado.
6. Realizar una serie de experimentos con la herramienta para poder comprobar los resultados arrojados.

7. Redactar una memoria que recoja todo el trabajo desarrollado, así como los manuales de instalación y de usuario.

## 1.4. Resultados esperados.

Al momento de finalización del presente proyecto, se pretende la consecución de los objetivos anteriormente expuestos. Los resultados del logro de estos objetivos han de ser:

- Estudio y conocimiento de diversas técnicas y alternativas para abordar el problema del análisis de emociones.
- La proyección de estos conocimientos y otras sugerencias para la obtención de un conjunto de léxicos que nos permitan el análisis de emociones en español.
- Generación de dichos recursos léxicos.
- Desarrollo de una herramienta que, a través de diferentes técnicas, permita el análisis de emociones.
- Creación de una memoria que recoja todo lo aprendido y realizado durante el desarrollo del proyecto.

## 1.5. Planificación temporal.

Para mostrar la planificación temporal que ha regido el presente proyecto, podemos ver los diferentes pasos a través de un Diagrama de Gantt, que nos permita ver el diferente tiempo que se le ha dedicado a cada una de las partes más importantes de este trabajo.

A continuación se expone en la tabla 1 cada una de las principales tareas que componen este proyecto y el tiempo que se le ha dedicado. También veremos el diagrama de Gantt consecuente obtenido de la tabla 1 en la ilustración 2:

| Planificación temporal | | | |
|---|---|---|---|
| **TAREA** | FECHA DE INICIO | DURACIÓN | FECHA FINAL |
| **Estudio de los lexicones** | 02/02/2018 | 28 | 02/03/2018 |
| **Estudio de generación de recursos** | 02/03/2018 | 7 | 09/03/2018 |
| **Generación del recurso léxico** | 09/03/2018 | 31 | 09/04/2018 |
| **Estudio de técnicas de clasificación** | 09/04/2018 | 15 | 24/04/2018 |
| **Diseño e implementación de la herramienta** | 24/04/2018 | 108 | 10/08/2018 |
| **Realización de pruebas** | 10/08/2018 | 73 | 22/10/2018 |
| **Creación de manuales de usuario y de instalación** | 15/10/2018 | 8 | 23/10/2018 |
| **Generación de la memoria** | 01/05/2018 | 175 | 23/10/2018 |

**Tabla 1. Temporalidad de las tareas**

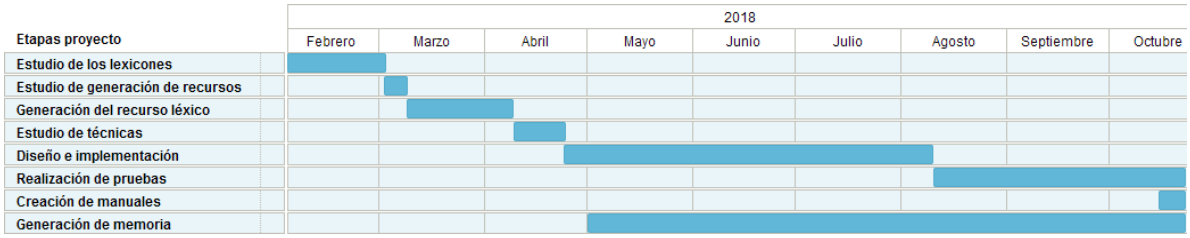



**Ilustración 2. Diagrama de Gantt**

Donde podemos apreciar la planificación seguida de una forma mucho más sencilla, además de observar qué partes han sido las que han conllevado mayor trabajo y en cuáles se ha invertido más tiempo en vista de su complejidad.

## 1.6. Estructura del proyecto.

En este punto, concluyendo el primer capítulo de introducción al proyecto, se mostrará la estructura de las diferentes partes de la memoria del presente trabajo, además de una breve explicación de los diferentes puntos. Este documento contiene

8 capítulos, incluyendo este punto de introducción, un apartado para la bibliografía utilizada y 4 anexos, dos para los manuales de instalación y de usuario y otros dos para los índices de tablas e ilustraciones:

- En el segundo capítulo se realizará un estudio de las emociones, objeto principal de este trabajo. Se abordará su significado desde distintas visiones para poder comprenderlo mejor de cara a realizar con conocimiento el resto del trabajo.
- El capítulo tres presenta un estudio de diferentes alternativas que se pueden utilizar para la generación del recurso, que es nuestro objetivo principal. Se analizarán distintas herramientas, sus características principales y se mostrará información sobre ellas.
- En el cuarto capítulo se explica y se desarrolla los diferentes métodos utilizados para la obtención del recurso léxico que más tarde será probado.
- El capítulo cinco presentará las diferentes técnicas existentes para la clasificación de emociones, con una explicación de sus principales características.
- En el sexto capítulo se expondrá los pasos seguidos para la creación de las herramientas analizadoras que se han desarrollado para probar los lexicones creados y el por qué se han elegido dichas opciones.
- El capítulo número siete mostrará los resultados de las distintas pruebas realizadas a través de las alternativas analizadoras en conjunción con los distintos lexicones, para dos tareas de clasificación diferentes que medirán su rendimiento.
- El último capítulo concluirá con los resultados y pensamientos obtenidos tras la consecución de todo el proceso que conlleva el presente trabajo. También se comentará la aplicación y ampliación en trabajos futuros.
- En el apartado destinado a la bibliografia se recogerán los distintos artículos y fuentes que se han utilizado como base para exponer, estudiar e investigar los diferentes componentes teóricos del presente trabajo.
- El manual de instalación incluirá las diferentes instalaciones que se han de realizar para poder usar la herramienta generada.

- Con el manual de usuario se podrá conocer cómo se ha de utilizar la herramienta creada para obtener el análisis deseado.
- El índice de ilustraciones recoge las diferentes imágenes que se han utilizado en el desarrollo del proyecto.
- Por último, el índice de tablas mostrará lo mismo que el índice anterior, pero con las tablas como objeto.

# 2. EMOCIONES. AFFECTIVE COMPUTING Y MODELOS DE EMOCIONES

## 2.1. El concepto de emoción

No es sencillo definir el concepto “emoción”, ya que resulta demasiado complejo. Tal y como dijeron Fehr y Russell (1984):

*“Everyone knows what an emotion is, until asked to give a definition”*

Cuya traducción vendría a decir:

“*Todo el mundo sabe lo que es una emoción hasta que se les pide una definición*”.

Teniendo en cuenta esta complejidad, para entender el significado del término, tanto de un modo general como de una manera más específica aplicada a nuestro problema, se va a proceder a analizar diferentes definiciones realizadas en distintos campos cuyo nexo común es la emoción.

Comenzando por una definición general, si acudimos a la definición de emoción otorgada por la Real Academia Española, nos encontramos con las siguientes acepciones [2]:

*1. f. Alteración del ánimo intensa y pasajera, agradable o penosa, que va acompañada de cierta conmoción somática.*

*2. f. Interés, generalmente expectante, con que se participa en algo que está ocurriendo.*

Donde vemos que la definición que más encaja con el sentido de la palabra “emoción” para este proyecto es la primera. Esta definición hace referencia a un amplio espectro de formas de sentir que quedan englobadas dentro del término “emoción”. Esta definición nos proporciona una idea general y básica acerca de lo que representa el concepto.

Por otro lado, tenemos la definición de Stanford Encyclopaedia of Philosophy de Ronald Da Sousa [3] [4]:

*"No aspect of our mental life is more important to the quality and meaning of our existence than emotions. They are what make life worth living, or sometimes ending."*

Que traducido al español vendría a decir:

"Ningún aspecto de nuestra vida mental es más importante para la calidad y significado de nuestra existencia que las emociones. Son las que hacen que merezca la pena vivir la vida, o a veces terminarla."

Como podemos ver, esta acepción es mucho más filosófica y más abstracta que la anterior, pero aunque el concepto de emoción converja en muchos campos de estudio, el análisis del mismo desde ellos es diferente [5]. Lo que podemos sacar en conclusión es que las emociones impregnan nuestras vidas en general y en particular, nuestras comunicaciones y relaciones [6].

Como se puede observar en [7], existe una tendencia a englobar dentro de la palabra "emoción" otras partes más específicas. Se puede ver entremezclado el significado de afecto y de emoción, siendo ambos conceptos muy generales. En [6] se recoge también, a través de una referencia a Scherer, la definición de cada una de las partes de lo que "afecto" engloba. La traducción e interpretación de las palabras en español vendría a ser:

- Emoción: episodio relativamente leve de respuesta a la evaluación de un acontecimiento externo o interno de gran importancia.
- Estado de ánimo: estado afectivo difuso, más pronunciado como cambio en la sensación subjetiva, de baja intensidad pero relativamente larga duración, a menudo sin causa aparente.
- Postura interpersonal: postura afectiva tomada hacia otra persona en una interacción específica, matizando el intercambio interpersonal en esa situación.
- Actitud: creencias, preferencias y predisposiciones hacia objetos o personas, relativamente duraderas y afectivamente influenciadas.

- Rasgos de personalidad: disposiciones estables de personalidad y tendencias de comportamiento cargadas emocionalmente típicas de una persona.

Tras estas definiciones, se puede apreciar cómo se ha englobado dentro de un término una amplitud de elementos que no tienen una conexión total entre sí y se ha podido discernir el significado concreto de emoción que va a ser el centro de este trabajo.

## 2.2. Análisis de sentimientos

Antes de proceder a comentar las principales teorías sobre las emociones, es conveniente enmarcar el uso de las emociones desde una perspectiva más cercana a nuestro campo de estudio. Si bien sabemos que en el estudio y el análisis de las emociones existe una confluencia de ciencias que trabajan con ellas, como la psicología, lingüística e ingeniería, nos vamos a centrar en la relación con la ingeniería, no desdeñando la labor de las otras partes y siendo conscientes de que gracias al análisis y al trabajo de las demás ciencias en cooperación se consigue la utilización en ingeniería de este elemento [8].

Como se ha mencionado en el punto anterior, las emociones impregnan nuestras vidas, nuestra existencia, por lo que, teniendo presente la creciente evolución en importancia de la tecnología en nuestras vidas diarias, si no se produjera una aceptación por parte de las tecnologías de nuestras emociones, no podríamos trabajar con ellas al nivel en el que ahora lo hacemos ni podrían ofrecernos las posibilidades que a día de hoy se nos presenta. Por ello, surge la necesidad de interconectar estos dos elementos tan dispares, y la primera persona en hablar de ello a través de la Computación Afectiva (Affective Computing) es Rosalind Picard en el año 1995 [9].

Una de las tareas principales de la Computación Afectiva es el tratamiento con emociones, pero lo cierto es que la Computación Afectiva es un campo muy amplio dentro del cual se encuadran diversas disciplinas.

Dentro de la Computación Afectiva nos encontramos con el Análisis de Sentimientos (Sentiment Analysis) [10], que a su vez es muy amplio y variado y que contiene las siguientes tareas que podemos ver en la ilustración 3:

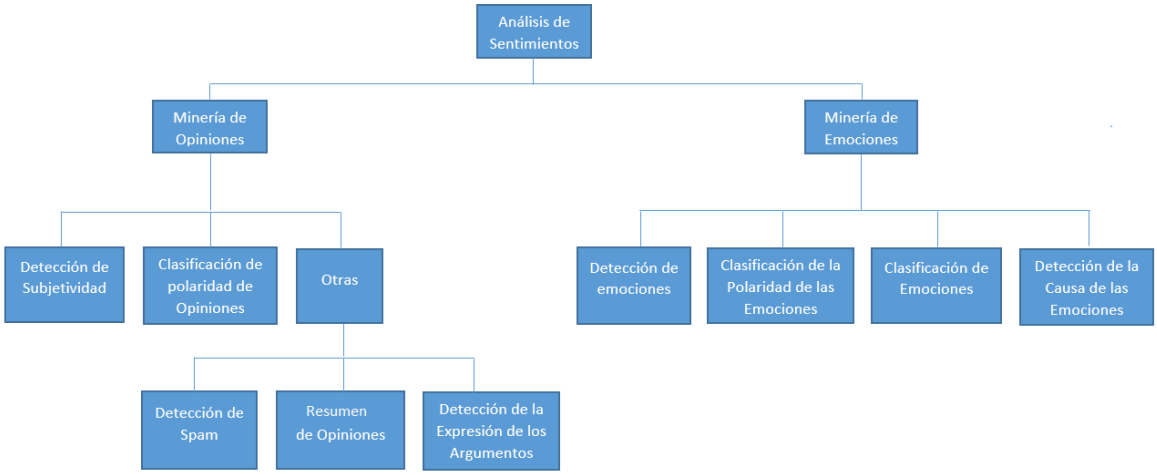



**Ilustración 3. Campos dentro del análisis de sentimientos**

Como vemos, existen dos principales campos de estudio: la minería de opiniones y la minería de emociones, pero debido a que el objeto del presente proyecto es la minería de emociones, únicamente vamos a analizar esta categoría.

Sin embargo, dentro de la misma se encuentran otras subtareas que tampoco son el objetivo del proyecto, que van a ser explicadas resumidamente, para poder centrarnos en la importante: la clasificación de las emociones.

- Detección de emociones: consiste en la detección de una determinada emoción en un texto, es decir, se basa en verificar la existencia de una emoción en un texto.
- Clasificación de la polaridad de las emociones: dado un texto, trata de determinar la polaridad de las emociones que contiene, dando por hecho que debe de existir alguna presente en él.
- Clasificación de emociones: consiste en la clasificación pormenorizada de las emociones que se encuentran presentes en un texto, siguiendo uno o varios conjuntos de emociones previamente definidos.

- Detección de la causa de la emoción: su cometido es la extracción de determinados tipos de factores para la adquisición de algunos tipos de emociones.

Si bien todas estas temáticas resultan fascinantes y se encuentran en su juventud de desarrollo, nos vamos a centrar en la clasificación de emociones, puesto que el propósito general del presente proyecto es el de generar un recurso léxico español que nos permita realizar clasificación de emociones.

Tal y como se entiende de la definición anterior de “clasificación de emociones”, se hace uso de un determinado conjunto de emociones para realizar este proceso de clasificación. En el punto siguiente, e hilando con el punto anterior, se van a exponer las principales teorías de modelos emocionales y algunos trabajos que son ejemplo de ellas.

## 2.3. Principales teorías sobre las emociones

Una vez estamos ubicados dentro de la Computación Afectiva para saber cómo afectan las emociones a la misma a través de la clasificación de emociones, hay que explicar con detalle las principales teorías sobre las emociones en lo que a modelos emocionales se refiere.

Existen mayoritariamente dos vertientes: el modelo básico de emociones y el modelo dimensional, si bien en algunos artículos se contempla otra tercera vertiente: la extendida. A continuación vamos a ver en profundidad cuál es la teoría que respalda cada una de estas vertientes y los principales ejemplos de cada una.

### 2.3.1. Teoría de las emociones discretas

La Teoría de las emociones discretas, Teoría de las emociones básicas o Estrategia de categorías emocionales [11] [12] [4] [13], divide las emociones en varios grupos y asigna una emoción dentro de su adecuada categoría discreta. Se argumenta que hay un conjunto limitado de emociones que una persona puede experimentar, como un conjunto de emociones modelo que son discretas. Estas

emociones básicas hacen referencia a aquellas que no tienen ninguna otra emoción como parte constituyente, o lo que es lo mismo, que desde un punto de vista biológico, las emociones de este conjunto surge de sistemas neuronales separados. Esta teoría es la más ampliamente extendida en la actualidad, debido a su sencillez y a que, pese a ella, otorga buenos resultados.

A continuación, podremos ver los principales estudios que siguen esta vertiente de los modelos de emociones.

- **Emociones básicas de Ekman**

Muchos de los trabajos realizados en esta vertiente de los modelos de emociones han sido llevados a cabo por el profesor Ekman. Su primera intervención en esta materia estaba basada en el reconocimiento de emociones básicas de expresiones faciales. Así, en el año 1974, Ekman fue uno de los primeros teóricos de emociones en sugerir que había un conjunto de emociones básicas que eran universalmente reconocidas. El compendio de emociones son las que se pueden observar en la tabla 2 y gráficamente en la ilustración 4:

| Enfado | Asco | Miedo | Alegría | Tristeza | Sorpresa |
|---|---|---|---|---|---|

**Tabla 2. Emociones básicas de Ekman**


**Ilustración 4. Emociones de Ekman**

Más tarde, Ekman creó una versión extendida de sus emociones básicas, quedando las emociones de la siguiente manera:

Enfado, Asco, Miedo, Alegría, Sorpresa, Diversión, Menosprecio, Contento, Vergüenza, Excitación, Culpa, Orgullo de Logro, Alivio, Satisfacción, Placer Sensorial, Humillación.

- **Emociones bipolares básicas de Plutchik**

Otro modelo distinto, también incluido dentro de esta categoría es el modelo de Plutchik. Este modelo define un conjunto de 8 emociones básicas bipolares, es decir, que están naturalmente emparejadas y que son opuestas, de modo que la activación de una emoción anularía la activación de su opuesta. Las emociones básicas de Plutchik se recogen en la tabla 3:

| Emoción | Opuesta |
|---|---|
| Alegría | Tristeza |
| Miedo | Enfado |
| Confianza | Aversión |
| Anticipación | Sorpresa |

**Tabla 3. Emociones básicas de Plutchik**

De estas ocho emociones básicas se puede realizar una composición para obtener emociones más complejas, y todo el conjunto de emociones es la llamada teoría de la Rueda de las Emociones de Plutchik, que se aprecia en la ilustración 5.

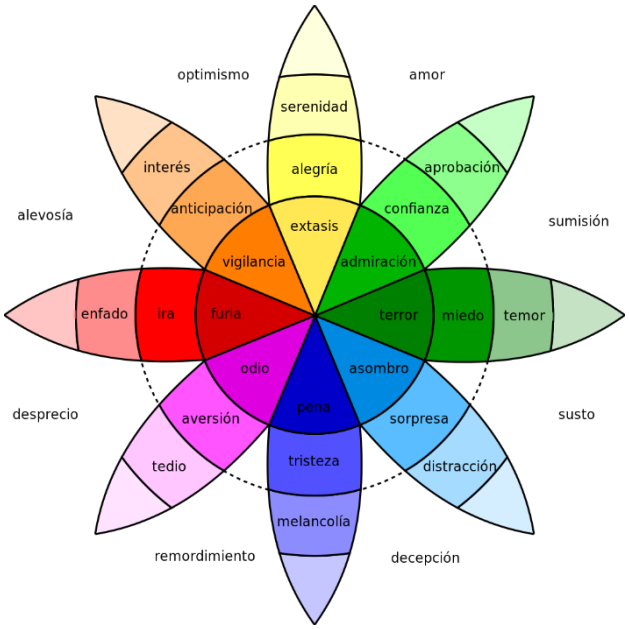



**Ilustración 5. Rueda de las emociones de Plutchik**

Los modelos de Ekman y Plutchik serán los utilizados para este proyecto, además de un conjunto de emociones que no tiene un origen como éstos, si no que se trata del sistema de puntuación de la página de noticias Rappler [14], puesto que es el conjunto en el que se basan en el lexicón Depeche Mood, que como se comentará más adelante será utilizado.

Además de estos dos modelos discretos, existen varios más, si bien son estas dos opciones las más extendidas. Otro modelos discreto:

- Modelo de Shaver [10]: Este modelo sigue una estructura de árbol en la que las emociones básicas son las ramas principales y cada rama tiene su propia categorización. Sería otra buena opción puesto que es bastante reducido en cuanto al conjunto de emociones que presenta, que podemos ver en la tabla 4:

| Enfado | Miedo | Alegría | Amor | Tristeza |
|---|---|---|---|---|

**Tabla 4. Emociones básicas de Shaver**

### 2.3.2. Modelos dimensionales de las emociones

La otra gran alternativa para afrontar el cómo formalizar las emociones es la teoría dimensional de las emociones [10] [12] [4] [13]. Puesto que no se va a utilizar

ningún conjunto de emociones fundamentado en este modelo, simplemente se expondrán sus principales consideraciones y se mostrarán brevemente algunos ejemplos.

Estas estrategias representan los efectos de una manera distinta al modelo anterior, a través de dimensiones, cuya forma de representación en el espacio se formaliza de manera que cada emoción ocupa una localización en el espacio, y cada una de las palabras se encuentran en estas dimensiones.

De un modo más relacionado con la neurociencia, lo que este tipo de estrategias quiere expresar es que no podemos entender las emociones como algo discreto, si no que conlleva la creación de un sistema común interconectado neurofisiológico que es responsable de todos los estados afectivos. Este sistema define unas emociones con respecto a una o más dimensiones que miden diferentes valores y magnitudes de las emociones, haciendo de ellas algo dinámico y continuo.

Los modelos más representativos de este tipo de estrategia son:

- Modelo Circumplejo de Afecto de Russell [12]: según este modelo, existe un espacio bidimensional que contiene las siguiente dimensiones:
    - Dimensión de valencia: indica cómo de agradable o desagradable es una emoción.
    - Dimensión de excitación: diferencia entre el concepto de activación y desactivación de una emoción.

- Modelo de Mehrabian [12]: esta estrategia habla de un modelo tridimensional cuyas dimensiones son:
    - Dimensión de valencia.
    - Dimensión de excitación.
    - Dimensión de dominancia: indica si un sujeto se siente en control de la situación o no.

### 2.3.3. Modelos extendidos

Existen además ciertos autores que consideran que existe una tercer modelo de emociones: el modelo extendido [11]. Esta categoría consiste en diversas estrategias que no están basadas puramente en ninguno de los dos modelos anteriores, sino que se valen de una combinación de ambos.

Algunos de los más reseñables, simplemente a fin de nombrarlos:

- Modelo de Ortony, Clore y Collins [11] [13]: en este estudio se centraron más en el evento que ha generado la emoción para determinar la emoción y evaluar la relación evento-emoción.
- Modelo de Cinco Factores [11]: en este caso, se utiliza más para identificar la tipología personal del escritor/lector.

Ambos presentan diversas diferencias con los otros modelos que hemos comentado anteriormente, pues no se centran en los mismos casos de estudio que se han visto, si no que van un paso más allá y toman en cuenta otros elementos que pueden mejorar el desempeño del análisis.

# 3. ESTUDIO Y ANÁLISIS DE RECURSOS

En este capítulo, se abordarán los diferentes recursos a utilizar para lograr la máxima del presente proyecto, esto es, lograr la generación de un lexicón en español. Antes de proceder al estudio de las diferentes alternativas que se nos presentan, y a fin de ubicar correctamente el contenido del capítulo, se deben de mencionar que:

- Existen tres estrategias para clasificación de emociones, principalmente. Estas tres estrategias son: estrategias basadas en lexicón, basadas en Machine Learning e híbridas. En el capítulo 5 se profundizará en cada una de ellas.
- A consecuencia de lo explicado en el punto anterior, todos los recursos que se expondrán son lexicones.

Una vez se han dejado claros los elementos anteriores, se va a proceder al análisis de diferentes opciones que resultan de gran interés.

La cantidad de recursos recogidos en la siguiente selección es reducida ya que el estudio de las emociones es una materia muy nueva aún, por lo que no existen tantos recursos como para otras vertientes, tal y como se comentó previamente en el capítulo de introducción. Aun así, existe una variedad suficiente como para poder escoger cuál de las soluciones se adecúa mejor al problema, pero como el problema que se presenta es muy amplio, se ha decidido realizar la generación de varios recursos en español, basados en diferentes herramientas que puedan proporcionar riqueza y que otorgan material para realizar pruebas de desempeño.

Así, y a la vista de las distintas alternativas, la selección escogida para los lexicones en español ha sido: NRC EmoLex [15], EmoSenticNet [16] y Depeche Mood [17]. Para responder al por qué se han seleccionado estos tres lexicones, se ha de ir paso por paso:

- En primer lugar, EmoLex es muy reconocido y presenta un conjunto de emociones bastante aceptable, que ni resulta corto ni excesivamente largo. Además, las emociones de Plutchik están ampliamente extendidas, e incluyen a otros conjuntos como son las emociones de Ekman.

Además, existía una versión en español para descarga, por lo que resultaba interesante ponerla a prueba.

- Con respecto a EmoSenticNet, su peculiaridad es que, a parte de presentar el conjunto de emociones de Ekman, un poco más reducido, contiene una gran cantidad de entradas compuestas por más de una palabra, lo que lo hace interesante como objeto de estudio. Además, al no estar en español, habría que crear un herramienta de traducción que nos permitiese obtener las palabras en español, lo cual también genera cierta atracción.
- Por último, en referencia a Depeche Mood, lo que lo diferencia de los otros es su diferente conjunto de emociones, que no está basado en ningún conjunto ya existente sino por las emociones utilizadas para puntuar las noticias de la ya comentada Rappler [14], y que además es más nuevo. También contiene un número muy alto de palabras, lo que lo hace un lexicón muy rico.

A continuación se mostrará información sobre cada uno de los recursos y, posteriormente, se comentarán en un último apartado las diferencias, similitudes, carencias y virtudes que presenta cada uno de ellos.

### 3.3. EmoLex (NRC lexicón)

Este recurso léxico fue fundado por Nacional Research Council de Canadá (NRC) [15], y fue creado a través de un servicio online de Amazon llamado Amazon Mechanical Turk, que requiere de “inteligencia humana” para desarrollar diferentes tareas. Además, se basa en el conjunto de emociones básicas de Plutchik, anteriormente descrito en el capítulo dos, puesto que consideraron que la cualidad bipolar de estas emociones resultaría de gran utilidad para el proyecto de creación de EmoLex. Respecto al número de entradas, este lexicón consta de 14182 conceptos.

El número de términos por emoción que presenta es:

- Enfado: 1247 términos
- Anticipación: 839 términos
- Asco: 1058 términos

- Miedo: 1476 términos
- Alegría: 688 términos
- Tristeza: 1191 términos
- Sorpresa: 534 términos
- Confianza: 1230 términos

Escogieron la plataforma de Amazon por varias razones, siendo las más reseñables la posibilidad de obtener una gran cantidad de anotaciones humanas de forma eficiente, no tener que realizar un gran desembolso y aún así obtener anotaciones de gran calidad. Sin embargo, y para que los resultados obtenidos fueran lo más satisfactorios posible, se debían definir las tareas cuidadosamente e incluso se tenían que realizar varias veces para asegurar que los datos recogidos de carácter aleatorio o erróneo fuesen debidamente descartados y rechazados.

A través de esta fórmula se genera el dataset EmoLex, cuyo tamaño es mayor que muchas otras herramientas utilizadas en este campo, como el lexicón WordNet Affect, entre otros.

Para la elaboración del compendio de palabras recogidas por este recurso léxico, se escogieron los verbos, adjetivos y nombres más frecuentes para el idioma inglés, obtenidos del diccionario Macquarie Thesaurus. Es reseñable mencionar que no solo trabaja con unigramas, sino que también se pueden encontrar un nada desdeñable número de bigramas, y también incluye términos del General Inquirer[1] [18] y del lexicón WordNet Affect[2] [15]. Así, de la unión de estas fuentes surgió un recurso con un número bastante elevado de términos.

Profundizando en cómo se realizaron las anotaciones, tras este proceso se esconde cierta complejidad, puesto que las palabras pueden tener diferentes sentidos y por tanto evocar diferentes emociones. Debido a esta complejidad surgen preguntas como qué sentido escoger y con qué granularidad representar los diferentes sentidos de una misma palabra. Un número elevado de sentidos diferentes conllevaría un exceso de información con el que no se podría generar un recurso consistente, pero

[1] http://www.wjh.harvard.edu/~inquirer/
[2] http://wndomains.fbk.eu/wnaffect.html

por otro lado trabajar solamente a nivel de palabra conllevaría una pérdida de calidad del sentido de la oración. De este modo, la solución que propusieron los creadores de este lexicón fue presentar primero un problema de elección en el que se pueda ver el sentido de la palabra en ese contexto, y se deba elegir entre cuatro opciones, y una vez realizado este proceso, contestar otra pregunta sobre cómo asociar el término en cuestión con cada una de las emociones utilizadas.

Estas asociaciones fueron puntuadas en una escala de 4, siendo (1) No, (2) Débilmente, (3) Moderadamente y (4) Fuertemente. Se eliminaron los valores anómalos y se adecuaron al sistema binario, quedando (1) y (2) dentro del valor 0 y (3) y (4) dentro del valor 1 [6].

### 3.4. EmoSenticNet

Este lexicón fue respaldado parcialmente por el Gobierno de India, México y de Reino Unido a través de becas. Consta de 13188 términos y el número de términos por emoción es de:

- Enfado: 829 términos
- Asco: 1159 términos
- Alegría: 9387 términos
- Tristeza: 1533 términos
- Sorpresa: 905 términos
- Miedo: 1199 términos

Propone una alternativa al léxico más utilizado hasta el momento de su creación, WordNet Affect, al igual que EmoLex (NRC). [16] Sin embargo, la estrategia del presente recurso es bastante diferente.

Hace uso de las etiquetas de emociones del WordNet Affect[3] utilizando los conceptos de SenticNet[4] a través de una estrategia de Machine Learning supervisada. La estrategia que apoya la creación de este recurso léxico es que las palabras que

[3] http://wndomains.fbk.eu/wnaffect.html
[4] https://sentic.net/downloads/

tienen características similares son semejantes en su uso y su significado y en particular están relacionadas con emociones parecidas. A continuación, se mostrará con un poco más de detalle cada uno de los recusos utilizados por este lexicón:

- SenticNet[5] [16]: es el lexicón objetivo y fuente de información sobre polaridad para poder realizar las mediciones de similaridad que se han comentado anteriormente. Por ello, se han utilizado todos los conceptos que tiene para la creación del EmoSenticNet. Este diccionario presenta valores de polaridad entre -1 y 1 para conceptos de una sola palabra y conceptos multipalabra. En la versión utilizada para la creación de este recurso léxico hay 5732 conceptos de los cuales 2690 entradas son multipalabra.
- Lista de emociones de WordNet Affect[6] [16]: como fuente de ejemplos usaron las listas de emociones de WordNet Affect que fueron utilizadas como dataset en el SensEval de 2007[7], que es una competición para sistemas de análisis semántico que luego derivó en SemEval [19]. Este conjunto de datos consta de seis listas de palabras que se corresponden con el conjunto de emociones básicas de Ekman, estudiado en el capítulo 2. Este elemento contiene 606 synsets, que son conjuntos de palabras que son sinónimos, y son la base de lexicón WordNet Affect.
- Dataset ISEAR[8] (International Survey of Emotion Antecedents and Reactions): este dataset se creó a través de una encuesta, tal y como su propio nombre indica. Dicha encuesta, que fue llevada a cabo en los años 90 en 37 países, consta de textos cortos llamados ‘statements’ (declaraciones), que se obtuvieron de aproximadamente 3000 encuestados que tenían que describir una situación en la que sentían una emoción particular.

  El dataset consta de 7666 ‘statements’, y para cada ‘statement’, el dataset contiene 40 parámetros numéricos o categóricos que proporcionan varios

[5] https://sentic.net/downloads/
[6] http://wndomains.fbk.eu/wnaffect.html
[7] http://web.eecs.umich.edu/~mihalcea/senseval/
[8] https://www.affective-sciences.org/research/materials-and-online-research/research-material/

tipos de información, como la edad del encuestado, la emoción sentida en la situación descrita, su intensidad, etc.

- Preprocesamiento y Tokenización: un conjunto de características que se usaron para la creación del lexicón estaban basadas en ocurrencias de conceptos en el dataset ISEAR. Para localizar estos conceptos en el texto, se emplearon herramientas de preprocesamiento del plug-in de texto de RapidMiner[9] para la lematización, además de WordNet lemmatizer. De los 5732 conceptos contenidos en SenticNet, 2729 fueron encontrados al menos una vez en ISEAR, tanto directamente como después de la lematización.
- Características usadas para clasificación: se extrajeron de ISEAR un conjunto de características estadísticas relacionadas con las ocurrencias y las co-ocurrencias de los conceptos de SenticNet (tratando los conceptos multipalabra como un solo token). Se utilizaron dos tipos de características: basadas en los parámetros que provee ISEAR y basadas en varias medidas de similitud entre conceptos.

  Con respecto a las basadas en los parámetros de ISEAR, se usaron 16 parámetros: información relacionada con el encuestado, datos generales sobre la emoción que sintió en la situación presentada, información fisiológica como sentir cambios de temperatura, información sobre su forma de comportarse…

  Sobre las basadas en medidas de similitud entre conceptos, se utilizaron medidas como la similitud basada en la puntuación de SenticNet, nueve medidas de similitud basadas en la distancia de WordNet, Punto de Información Mutua, Afinidad Emocional…

Finalmente, para el proceso de clasificación se realiza una tarea de categorización con respecto a las seis emociones contempladas en este lexicón, es decir, a cada uno de los conceptos con los que trabajan se les asigna una de estas seis emociones.

[9] http://rapid-i.com/content/view/181/190

### 3.5. Depeche Mood

Depeche Mood es un lexicón cuya creación ha sido apoyada parcialmente por el proyecto Trento RISE PerTe, y se trata del recurso más nuevo de los tres que vamos a emplear [17]. Además, fue creado de una manera novedosa y totalmente automatizada a través de crowdsourcing en la página web de noticias Rappler [14]. Este lexicón consta de un total de 37772 entradas y las emociones básicas con las que trabaja son: Susto, Diversión, Enfado, Molestia, Indiferencia, Felicidad, Inspiración y Tristeza. La distribución de términos según cada una de estas emociones es:

- Susto: 25860 términos
- Diversión: 31952 términos
- Enfado: 28549 términos
- Molestia: 29754 términos
- Indiferencia: 29192 términos
- Felicidad: 35326 términos
- Inspiración: 32319 términos
- Tristeza: 29774 términos

Concretamente, para su desarrollo se utilizaron los artículos etiquetados de noticias de la página web Rappler, anteriormente referenciada, obteniendo un dataset con decenas de millones de palabras en más de 20000 documentos, teniendo más de 500 palabras por documento de media. Para categorizar dichos documentos se utiliza un medidor del estado anímico, original del sitio web, que ofrece a los lectores la oportunidad de expresar las emociones que la noticia que están leyendo les provoca. De este modo, se pueden obtener datos etiquetados por los usuarios que sirven para el proceso de creación, sin que sean realmente conscientes de que se utilizan para eso. A través de esto, se crean matrices documento-emoción.

Para la creación del lexicón, se construyó una matriz palabra-emoción a través de una estrategia de semánticas compuestas desde la matriz documento-emoción. Los pasos de este proceso fueron los siguientes:

- Primero se realizó la lematización y análisis de la categoría gramatical (Part Of Speech) para todos los documentos, almacenando solo aquellos que estuviesen presentes en el lexicón WordNet[10] en la matriz documento-emoción.
- Posteriormente, se calcularon las matrices palabra-documento utilizando las frecuencias obtenidas sin procesar, frecuencias normalizadas y tf-idf.
- Una vez obtenido esto, se multiplicaron las matrices documento-emoción y palabra-documento para obtener una matriz palabra-emoción, para poder unir palabras con emociones sumando los productos de los pesos de cada palabra con el peso de las emociones en cada documento.

A través de esta matriz palabra-emoción se crearon las tres versiones diferentes del lexicón: sin normalizar, normalizada y tf-idf.

Con respecto al conjunto de emociones, a continuación se mostrarán en la tabla 5 las diferentes emociones y su equivalencia en las emociones utilizadas en SemEval, que como se comentó en la explicación del lexicón anterior, es una competición de sistemas para análisis semántico que anteriormente se llamaba SensEval [19]:

| SemEval | Rappler |
|---|---|
| Miedo | Asustado |
| Enfado | Enfadado |
| Alegría | Feliz |
| Tristeza | Triste |
| Sorpresa | Inspirado |
| - | Molesto |
| - | Divertido |
| - | Indiferente |

**Tabla 5. Traslado de las emociones de Rappler a las emociones de SemEval**

La tabla 5 ha sido obtenida de [17], donde podemos apreciar cuáles serían las coincidencias entre las emociones de SemEval y las de Depeche Mood. Como podemos ver, no existe equivalencia para todas las emociones, dato que resulta de mucha utilidad puesto que la fase de experimentación va a ser realizada con datasets

[10] http://wndomains.fbk.eu/wnaffect.html

de SemEval 2018[11], y esta aproximación nos va a permitir poder utilizar las emociones de Depeche Mood para la tarea de clasificación.

### 3.6. Comparación de los lexicones

Tras explicar brevemente cómo se creó cada uno de los recursos léxicos que se van a utilizar, el siguiente paso consiste en realizar una comparación de ellos, y de sus resultados en inglés [20].

| Lexicón | EmoLex | EmoSenticNet | DepecheMood |
|---|---|---|---|
| **Origen** | Crowdsourcing | - | Crowdsourcing |
| **Pago de anotaciones** | Sí | No anotadas | No |
| **Emociones** | 8 (Plutchik) | 6 (Ekman) | 8 (Rappler) |
| **Número de elementos** | 14182 | 13188 | 37772 |
| **Multipalabra** | Sí | Sí | No |

**Tabla 6. Comparación de diferencias entre los distintos lexicones**

Respecto a sus diferencias, podemos decir, como podemos ver en la tabla 6 que:

- Mientras que EmoLex y Depeche Mood han sido obtenidos a través de crowdsourcing, el método de generación de EmoSenticNet es más tradicional. Sin embargo, dentro del método de crowdsourcing de las otras dos propuestas existe una diferencia principal: las personas que realizaron el proceso de anotación de EmoLex fueron retribuídas por su trabajo a través de Amazon Mechanical Turk y eran plenamente conscientes de ello, mientras que las personas que hicieron la misma labor pero para Depeche Mood no recibieron ningún tipo de pago y no eran conscientes de para qué se utilizaron sus anotaciones.
- Otra desemejanza más llamativa es los diferentes conjuntos de emociones que cada uno de los elementos presenta. Este factor otorga un punto de vista alternativo para cada uno de los lexicones pero también

[11] https://competitions.codalab.org/competitions/17751

genera el problema de no poder realizar comparaciones totales de desempeño entre ellos tres.

- Con respecto a la cantidad de elementos, si bien a priori EmoLex y EmoSenticNet tienen un número similar, no todos los valores de EmoLex aportan un valor real, por lo que su verdadero número de items que proporcionan información es menor al que tienen. Por otro lado, Depeche Mood tiene una gran cantidad de elementos, superando con creces, incluso triplicando el número de entradas de los otros dos lexicones.
- En relación al número de palabras presentes en cada fila de cada lexicón, cabe destacar que, mientras que en Depeche Mood y en EmoLex la gran mayoría son una palabra, en EmoSenticNet existe una alta cantidad de dos palabras y una inferior de tres o más palabras, pero mayor que en los otros dos elementos.

Una vez expuestas las anteriores disensiones entre recursos, se va a proceder al estudio de la diferencia de resultados. Para ello, en [20] realizaron un cuestionario en forma de mensajes/textos de Facebook o noticias que los lectores, sin saberlo, anotaban, permitiendo el análisis con las diferentes implementaciones de los lexicones que desarrollaron, y utilizando diversas técnicas de Procesamiento del Lenguaje Natural como lematización e identificación de la categoría gramatical, y diferentes heurísticas para el tratamiento de los adverbios de cantidad y de los negadores.

Como conjunto de emociones, utilizaron las emociones comunes de los recursos excepto ‘Sopresa’, es decir, utilizaron ‘Miedo’, ‘Enfado’, ‘Alegría’ y ‘Tristeza’, y como métricas se apoyaron en la medida de Precisión, de Exhaustividad y el valor F. Los resultados que obtuvieron se ven reflejados en la tabla 7:

| | NRC | | | DepecheMood | | | EmosenticNet | | |
|---|---|---|---|---|---|---|---|---|---|
| | P | R | F | P | R | F | P | R | F |
| Negative | **84%** | **86%** | **0.85** | 83% | 70% | 0.76 | 88% | 20% | 0.32 |
| Anger | **37%** | **68%** | **0.48** | 50% | 28% | 0.36 | 29% | 7% | 0.11 |
| Fear | **48%** | **34%** | **0.40** | 41% | 12% | 0.18 | 56% | 1% | 0.03 |
| Sadness | 53% | 27% | 0.36 | **34%** | **56%** | **0.42** | 56% | 23% | 0.33 |
| Positive (joy) | **56%** | **52%** | **0.54** | 39% | 57% | 0.46 | 28% | 92% | 0.42 |

**Tabla 7. Comparación de resultados entre lexicones**

Los resultados de esta tabla han sido obtenidos de [20], donde se puede apreciar, según este estudio, si miramos el resultado del valor F, que los que mejores resultados arrojan son EmoLex(NRC) y luego DepecheMood, si bien es notorio que EmoLex (NRC), con diferencia, es el que mejores resultados suele ofrecer. Con respecto al significado de Negativo y Positivo, se engoblan las emociones ‘Enfado’, ‘Miedo’ y ‘Tristeza’ dentro de Negativo, para estudiar si cada lexicón es mejor para analizar emociones positivas o negativas.

Si sólo miramos el valor de Precisión, la situación es un poco más difusa, puesto que para cada emoción hay un lexicón diferente que ofrece el mejor resultado, aunque el mejor sería EmoSenticNet puesto que es el mejor para dos emociones.

Por último, si sólo miráramos la exhaustividad, el que mejores resultados ofrece es EmoLex (NRC).

Como nexo de estas dos medidas, tenemos el valor F, que el que mejores resultados devuelve es EmoLex, teniendo en cuenta que la fórmula del valor F implica Precisión y Exhaustividad.

Teniendo en cuenta estos resultados, en pos de un análisis en inglés, sería interesante, haciendo una combinación de los tres recursos, ordenar los resultados según las emociones en las que cada uno arroja mejores resultados.

# 4. GENERACIÓN DEL LEXICÓN

En este capítulo se van a estudiar las distintas técnicas empleadas para la generación del recuros léxico en español. En este caso, se han creado 5 lexicones distintos (uno con EmoLex, otro con EmoSenticNet y tres con DepecheMood) para los procesos de análisis de emociones.

A continuación, se van a explicar las estrategias seguidas para la creación de las tres alternativas, ya que para cada uno de ellos ha habido que realizar un proceso de tratamiento distinto.

## 4.3. Herramientas de traducción utilizadas

En primer lugar cabe mencionar que para la generación de los léxicos se han utilizado herramientas de traducción automática, puesto que aunque EmoLex se puede encontrar en español traducido descargado desde la fuente oficial, las otras alternativas no presentaban traducciones oficiales, por lo que se han tenido que generar con traducción automática.

Las herramientas de traducción automática serán explicadas en el correspondiente apartado pero los motores de traducción que se han utilizado son:

- Traductor de Google: el traductor más famoso, bastante potente y que ofrece presumiblemente fidedignas traducciones, si bien es cierto que la calidad de las mismas en ocasiones no es la mejor. Aun así, se trata de un buen traductor y ofrece una gran cantidad de idiomas para traducir.
- Traductor DeepL: es un traductor bastante nuevo, pero que devuelve unos resultados de traducción muy acertados y de una gran calidad utilizando para ello redes neuronales y la web Linguee. Con respecto a la calidad de traducciones se podría decir que es mejor que la de Google pero tiene limitaciones, aún así es capaz de devolver traducciones muy naturales.
- Traductor de IBM: este motor también genera unos resultados muy buenos, con una calidad muy similar a los producidos por DeepL, con la historia y el prestigio de IBM.

En comparación, los que mejores traducciones ofrecen en cuanto a la calidad de traducción y a la naturalidad son probablemente DeepL e IBM, si bien los resultados de Google tampoco se podrían considerar malos.

## 4.4. Tratamiento de EmoLex (NRC)

Para este lexicón se ha utilizado la versión ya existente en español, si bien cabe destacar que esta versión es una versión traducida del original a través del traductor de Google, por lo que no es una traducción totalmente fidedigna.

Otra de las características destacables es que hay muchas palabras que no tienen ningún tipo de carga emocional, es decir, hay muchas palabras para las que todos los valores de cada una de las emociones es cero. Teniendo esto en cuenta, se ha procedido a una limpieza del mismo para eliminar estas entradas que no aportan nada.

| Spanish Translation | Anger | Anticipation | Disgust | Fear | Joy | Sadness | Surprise | Trust |
|---|---|---|---|---|---|---|---|---|
| detrás | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ábaco | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| abandonar | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| abandonado | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| abandono | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| disminuir | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| disminución | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Tabla 8. Ejemplos de valores para EmoLex**

Como podemos apreciar en la tabla 8, algunos valores sí que presentan valores distintos de cero, pero por ejemplo “detrás”, que aparece en el lexicón, no aporta ningún tipo de valor al mismo ya que todos sus valores son iguales a cero.

Si el conjunto de palabras que forman este lexicón son alrededor de 14000 palabras, más de la mitad no tienen ningún tipo de carga emocional. Por ello, tras el proceso de eliminación nos quedamos con en torno a 3500 palabras, tal y como se puede apreciar en la tabla 9, lo que resulta impactante por la diferencia de tamaños.

Por tanto podemos concluir que el tratamiento realizado para éste léxico ha sido el que menos trabajo ha dado. Simplemente, se ha realizado la descarga de este lexicón a través del enlace oficial[12] y se ha procedido a su procesamiento, ya que

[12] http://sentiment.nrc.ca/lexicons-for-research/NRC-Emotion-Lexicon.zip

presentaba formato de hoja de excel y es mucho más sencillo trabajar con formato .txt.

| Versión | EmoLex |
|---|---|
| **Original (Inglés)** | 14182 |
| **Traducida** | 11356 |
| **Procesada** | 3464 |

**Tabla 9. Número de palabras de EmoLex**

## 4.5. Tratamiento de EmoSenticNet

Para el siguiente lexicon sí se ha tenido que realizar traducción puesto que no existía una versión en español de la fuente oficial[13]. Para obtener la versión en español se ha realizado un proceso de traducción automática utilizando los tres traductores explicados en el apartado 4.1. Además, al encontrarse en formato excel, se ha tenido que realizar el tratamiento y se ha pasado a archivo de texto para facilitar las lecturas futuras.

Respecto al método de traducción, el proceso seguido ha sido el siguiente, para todas y cada una de las filas del lexicón:

- En primer lugar se intenta traducir con el traductor de Google, ya que es el más ampliamente utilizado y el más rápido de utilizar.
- Si no devuelve traducción, ya que a veces la herramienta utilizada no devolvía una posible traducción, se pasa al siguiente traductor, que es DeepL.
- Si con DeepL ocurre lo mismo, se prueba por último con IBM.
- Si logramos una traducción, se añade a una lista de palabras similares.
- Por último, se busca la palabra en inglés en los synsets de WordNet que están presentes en la librería de Python NLTK[14], y se devuelven sinónimos en español. Si hay sinónimos, se almacenan (junto con la traducción, si la hay) con las mismas puntaciones emocionales que la palabra original en el nuevo lexicón.

[13] https://www.gelbukh.com/emosenticnet/emosenticnet.xls
[14] https://www.nltk.org/

Tras esto, se realiza un proceso de eliminación de palabra duplicadas por mayúsculas y minúsculas y se eliminan tildes y caracteres poco comunes. En el caso de producirse la situación de que una palabra duplicada tuviese diferentes valores, simplemente se almacenaría los valores de la primera de las palabras que se lee, porque al introducirse la siguiente, como la anterior ya existiría en el documento, no se añadiría. Los resultados obtenidos, con respecto a la variación del número de palabras es el que se muestra en la tabla 10:

| Versión | EmoSenticNet |
|---|---|
| **Original (Inglés)** | 13188 |
| **Traducida** | 15660 |
| **Procesada** | 14058 |

**Tabla 10. Número de palabras del lexicón EmoSenticNet**

## 4.6. Generación de Depeche Mood

Por último, en el caso de Depeche Mood también se ha tenido que realizar traducción de las palabras, puesto que tampoco existía una versión en español reconocida.

Con respecto al tratamiento de lectura, al ser un documento de texto, ha sido un proceso más sencillo que con los dos lexicones anteriores. Pero una peculiaridad es que tras cada palabra hay una almohadilla y después una letra (“a”, “n”, o “v” dependiendo de si es adjetivo, nombre o verbo). Por ese motivo, se ha decidido dividir el lexicón original en tres lexicones distintos, uno para cada tipo de palabra. De hecho, es este cambio lo que afecta al proceso de traducción anteriormente descrito. Por ello, se ha recorrido cada uno de estos tres nuevos lexicones y se ha ido palabra por palabra siguiendo estos pasos:

- Se ha intentado traducir la palabra con Google. Si la palabra no era de la misma categoría gramatical que la palabra original, se descartaba la traducción.
- En caso de no haber traducción o haber sido rechazada, se intenta con DeepL, y se repite la comprobación anterior.
- Si de nuevo no obtenemos una traducción, se repite de nuevo con IBM.

- Si hemos obtenido una traducción válida se almacena, y pasamos la palabra original para obtener los sinónimos en español, asegurando que estos sinónimos son del mismo tipo de palabra que la palabra original.
- Una vez obtenida esta lista de palabras, se les asigna en el nuevo lexicón las emociones de la palabra original.

Tras esto, se ha realizado una normalización de las palabras para evitar duplicidades, y poder hacer que el máximo posible de elementos del lexicón sean útiles.

Tras esto, se realiza un proceso de eliminación de palabra duplicadas por mayúsculas y minúsculas y se eliminan tildes y caracteres poco comunes. En el caso de producirse la situación de que una palabra duplicada tuviese diferentes valores, se trataría igual a como se ha comentado en la sección de EmoSenticNet. El resultado, con respecto al número de palabras, se ve reflejado en la tabla 11:

| Depeche Mood | | | |
|---|---|---|---|
| **Versión** | **Adjetivos** | **Nombres** | **Verbos** |
| **Original (Inglés)** | 9066 | 21549 | 5348 |
| **Traducida** | 8638 | 24626 | 5559 |
| **Procesada** | 7921 | 21055 | 4956 |

**Tabla 11. Número de palabras de Depeche Mood**

# 5. ESTUDIO DE TÉCNICAS PARA LA CLASIFICACIÓN DE EMOCIONES

Con respecto a las distintas técnicas que se pueden emplear para el proceso de clasificación, se siguen mayoritariamente dos estrategias principales y una tercera híbrida de ambas.

Estas dos principales estrategias son: técnicas basadas en recursos léxicos y técnicas basadas en Machine Learning. Mientras que el primer tipo de técnicas dependen de fuentes léxicas, las basadas en Machine Learning aplican un algoritmo de Machine Learning que realice este proceso de detección de emociones.

A continuación se verán con más detalle estas dos técnicas y sus características y se explicará superficialmente la híbrida.

## 5.2. Estrategias basadas en fuentes léxicas

Las estrategias basadas en recursos léxicos son aquellas que dependen de un lexicón con el que poder realizar la tarea de detectar emociones. Existen dos estrategias principales para generar lexicones, aparte de la manual [21]:

- Estrategia basada en diccionario: esta estrategia trata de encontrar unas palabras ‘semilla’ que se puedan analizar, en este caso, emocionalmente. Después, se realiza una búsqueda en el diccionario para poder encontrar sinónimos y antónimos que poder anotar a través de la carga emocional de la palabra semilla.
- Estrategia basada en corpus: esta estrategia comienza con un lista de palabras semilla, a través de las cuales encuentra otras palabras en un corpus que puedan ayudar en la búsqueda de palabras con orientaciones específicas del contexto. Suele usar métodos estadísticos o semánticos.

## 5.3. Estrategias de Machine Learning

Con respecto a las estrategias basadas en Machine Learning, hay que mencionar que este es un campo que se ocupa de la construcción y estudio de diferentes algoritmos que puedan realizar la tarea de aprendizaje de los datos.

Estos algoritmos se basan en la construcción de un modelo que tiene unas determinadas variables de entrada que, utilizándolas, realiza predicciones o toma decisiones, en vez de seguir sólo instrucciones. Por tanto, su aplicación en la detección de emociones es el empleo de este aprendizaje para poder realizar la clasificación.

Existen dos tipos de vertientes [10] [12]:

- Aprendizaje supervisado: se usan para el aprendizaje datos previamente etiquetados, a los que se conoce como datos de entrenamiento. Los diferentes algoritmos los analizan para más tarde poder realizar la inferencia de una función que permite la clasificación de nuevos elementos.

  Por conjunto de datos etiquetados, o corpus, entendemos a un conjunto de ejemplos a los que se les ha asociado unas puntuaciones con respecto a la intensidad o existencia de una emoción en cada uno de los ejemplos.

  La principal desventaja es precisamente este proceso de anotación de los datos, que puede resultar tedioso. Sin embargo, existen procesos automáticos para la generación de etiquetas para corpus que pretenden subsanar este inconveniente.

- Aprendizaje no supervisado: se basan en la búsqueda de estructuras ocultas en un conjunto de datos no etiquetados a través de las cuales construir los modelos de clasificación de emociones.

## 5.4. Estrategia híbrida

El método híbrido [10] trata de la combinación de las dos estrategias anteriores, si bien se abre un abanico de posibilidades en el cómo y en qué medidas usar una u otra. Principalmente, la utilidad que reside en este método es que intenta equilibrar las desventajas de cada uno de los métodos anteriores combinándolos e intentando llevarlos a los mejores resultados. De hecho, en un gran número de ocasiones, esta es la táctica seguida para el mejor rendimiento.

# 6. ESTRATEGIA DE CREACIÓN DEL ANALIZADOR DE EMOCIONES

En este capítulo se va a proceder a la explicación y desarrollo del método seguido para la creación de los diferentes analizadores que se han generado para poder probar los recursos léxicos que han sido producidos.

Para poder probar la efectividad de los lexicones, y poder poner a prueba su funcionamiento con diferentes alternativas, se ha decidido crear tres analizadores diferentes que permitan clasificar las emociones según cada lexicón y atendiendo a diferentes estrategias con las que han sido abordados. Pero antes de comentar las diferencias que existen entre cada uno de ellos, se han de presentar los métodos comunes utilizados en las distintas heurísticas de clasificación.

Se trata de un conjunto de tratamientos y de funciones que permitan maximizar los resultados y valorarlos de una manera más acorde, que pueda proporcionar un mejor rendimiento y que haga por tanto que no se base la tarea del analizador en una simple suma de valores.

Así, los métodos comunes empleados son:

- Lematización y tokenización de los textos: permiten realizar el tratamiento de los elementos y posibilitan el análisis de las palabras. A través de la lematización podemos transformar una forma verbal en su correspondiente verbo en infinitivo y así poder realizar de un modo más efectivo su posterior procesamiento, sin dividir la tarea en cada expresión y por tanto perder el valor.
- Eliminación de tildes y de caracteres extraños: consiste en la supresión de estos elementos y en la consecuente transformación de las palabras que los contienen a palabras normalizadas. Puesto que los términos de los lexicones están normalizados, para asegurar que podemos encontrar los valores se ha procedido a realizar este tratamiento.
- Eliminación de las palabras vacías: se trata de la eliminación de aquellas palabras que estén dentro de la lista de palabras vacías porque no aportan

ningún valor al texto, por lo que sería inútil realizar su búsqueda en los lexicones y buscar su carga emocional.

- Clasificación: con respecto a la heurística que se va a utilizar para la clasificación, se ha de comentar que es muy simple. Se trata de la suma de las puntuaciones de las palabras (o bigramas o estructuras) en los diferentes lexicones empleados.
- Tratamiento de los adverbios modificadores: esta tarea se ha dividido en tres subtareas: tratamiento de la negación, de los adverbios modificadores que aumentan el valor, y los adverbios modificadores que restan valor:
    - Negación: se ha creado una lista de negadores en español que posibilite la búsqueda de estos términos en los textos. El cómo se ha realizado depende del tipo de analizador, pero su función ha sido la de eliminar la emoción de la siguiente palabra cargada de emoción que se encontrase en el caso de los lexicones EmoSenticNet y Depeche Mood, y en el caso de EmoLex, realizar el incremento de la emoción bipolar que se opone a la emoción negada.
    - Adverbios modificadores que aumentan el valor: son aquellos que magnifican la intensidad de la palabra a la que acompañan. Al igual que para los negadores, también se ha creado una lista con las palabras principales en esta categoría. Un ejemplo:
        - Este artículo me parece caro.
        - Este artículo me parece extremadamente caro.

      El valor de “caro” en las dos oraciones no es el mismo, en la segunda oración se ve intensificado. Por eso, en caso de encontrarse una palabra perteneciente a esta lista de palabras intensificadoras, el valor de la emoción de la palabra a la que acompañan se verá aumentado al ser multiplicado dicho valor por 1,3.

    - Adverbios modificadores que disminuyen el valor: en este caso se produce la misma situación que con los adverbios del apartado

anterior, pero al contrario, es decir, disminuyen la intensidad de la palabra a la que acompañan. Así tenemos:

- La tostada está quemada.
- La tostada está ligeramente quemada.

De nuevo, no tiene la misma intensidad quemada en las dos oraciones. Por eso, en el caso de localizar en el texto alguna palabra que se encuentre en la lista de adverbios modificadores que disminuyen el valor que se ha creado con anterioridad, el valor de las emociones de la palabra por la que esté acompañada este adverbio se verá disminuido, ya que se multiplicará sus valores de emoción por 0,4.

o Tratamiento especial para verbos y adjetivos con respecto a los nombres y los adverbios: para cada oración del texto que se facilite, se realiza un análisis con PoS tagger, es decir, un análisis de la categoría gramatical, para saber la de cada una de las palabras que la forman. De este modo, si se encuentra que una palabra está en el lexicón utilizado y que esta palabra es un verbo o un adjetivo, se les aumentará el valor de las emociones que tenga asociada dicha palabra al mutiplicar dicho valor por 1,5.

Si se trata, por el contrario, de un nombre o un adverbio diferente a los anteriormente mencionados, su valor se verá menguado ligeramente al multiplicar dicho valor por 0,9.

Una vez explicados los diferentes métodos seguidos para todos los analizadores, se puede pasar a mostrar las tres herramientas creadas y la estrategia que hay detrás de cada una de ellas.

## 6.3. Análisis simple

Para este primer ejemplo, se ha realizado un análisis simple, esto es, un análisis palabra por palabra. Tal y como se comentó en el capítulo de explicación de los

lexicones, algunos de ellos contienen bigramas, o si bien en su concepción inicial no los tenían, en su traducción al español los han obtenido.

Sin embargo, la estrategia detrás de este analizador no contempla la posibilidad de poder analizar conjuntamente grupos de palabras, si no que trata palabra por palabra. Eso hace que sea el más simple de los tres, puesto que su única labor más allá de la de las funciones explicadas con anterioridad es la de separar por palabras la oración e ir tratándolas una a una.

A continuación, a fin de entender mejor los pasos seguidos por el analizador, se va a mostrar un diagrama aclaratorio en la ilustración 6 y un ejemplo:

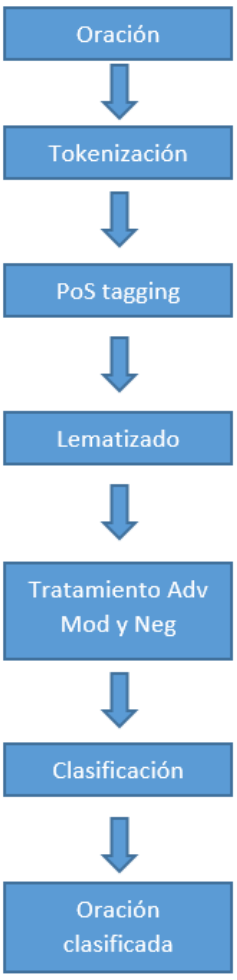



**Ilustración 6. Diagrama del análisis simple**

Donde se pueden apreciar claramente los pasos: en primer lugar se tokeniza y se realiza PoS tagging. Una vez hecho esto, se lematizan las palabras, se eliminan las tildes y se hace tratamiento especial para los modificadores y los negadores.

Se intenta realizar la clasificación de cada una de las palabras, teniendo en cuenta los modificadores y negadores, y para obtener la oración completa clasificada se emplea la suma de los resultados para cada palabra.

Como ejemplo vamos a utilizar el lexicón EmoLex (NRC). En el programa creado para este analizador, se va a escribir la siguiente oración: “Esto es muy absurdo”. Contiene “muy”, que es un modificador que aumenta la intensidad, y “absurdo”, cuya puntuación en el lexicón es la que se muestra en la ilustración 7:

```
absurdo: [0.0, 0.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0]
```

**Ilustración 7. Resultados de emociones para 'absurdo'**

Que presenta valor 1 para tristeza. Ya teniendo esto en cuenta, ejecutamos el analizador con esa oración. El resultado obtenido es el mostrado en la ilustración 8:

```
Introduzca el texto que desea analizar:
Esto es muy absurdo
¿Con qué lexicón quiere utilizarlo? Elija la opción deseada:
0. NRC
1. EmoSenticNet
2. Depeche Mood
0
muy
absurdo [0, 0, 0, 0, 0, 1.9500000000000002, 0, 0]
Resultados obtenidos con NRC:
[0, 0, 0, 0, 0, 1.9500000000000002, 0, 0]
Enfado: 0
Anticipación: 0
Desagrado: 0
Miedo: 0
Alegría: 0
Tristeza: 1.9500000000000002
Sorpresa: 0
Confianza: 0
```

**Ilustración 8. Resultados de análisis simple**

Como se puede observar, detecta “absurdo”, y como está precedido por “muy”, se intensifica su valor, pasando de 1 a 1,95.

## 6.4. Análisis a través de estructuras gramaticales

Este analizador es el más complejo de los tres. La idea detrás de él reside en que, si estudiamos las categorías gramaticales que existen en una oración, y las pasamos por el lexicón, quizá podamos obtener valores emocionales de ese conjunto de palabras más fidedignos que los que podrían generar las palabras por separado.

El proceso seguido para ello ha sido:

- Encontrar las estructuras gramaticales.
- Eliminar las palabras vacías.
- Compararlas con los lexicones.
- Si están, se utilizan esos valores emocionales.
- Si no están, se separa la estructura y se analiza cada palabra individualmente a través del analizador simple.

El procedimiento es más complejo, pero esa es la idea general. A raíz de contemplar el tratamiento de estas entradas del lexicón compuestas por más de una palabra, también se han estudiado casos especiales relacionados con las negaciones y con los adverbios modificadores, de manera que si existía la presencia de ellos en la estructura gramatical, se procedía a eliminar esta palabra del conjunto, buscar este nuevo conjunto entre el lexicón y, en el caso de ser encontrado, aplicar sobre él los valores de las modificaciones que suponga el adverbio previamente eliminado.

Así, se asegura que podremos encontrar el análisis anterior, pero con la posibilidad de poder obtener un análisis más adecuado para conjuntos encontrados.

Al estar basada en estructuras gramaticales, el número de elementos puede variar, si bien en la mayoría de las ocasiones es de un solo elemento, algunas veces de dos y muy esporádicamente de tres o más.

Este análisis de estructuras solo se realiza con los lexicones EmoLex y EmoSenticNet, puesto que el lexicón Depeche Mood tiene una alta cantidad de entradas con un solo elemento, que hacen que someterlo a este análisis más tedioso y más lento que el primero sea probablemente innecesario.

A continuación en la ilustración 9 se va a mostrar a través de un diagrama el funcionamiento de este analizador de una manera superficial:

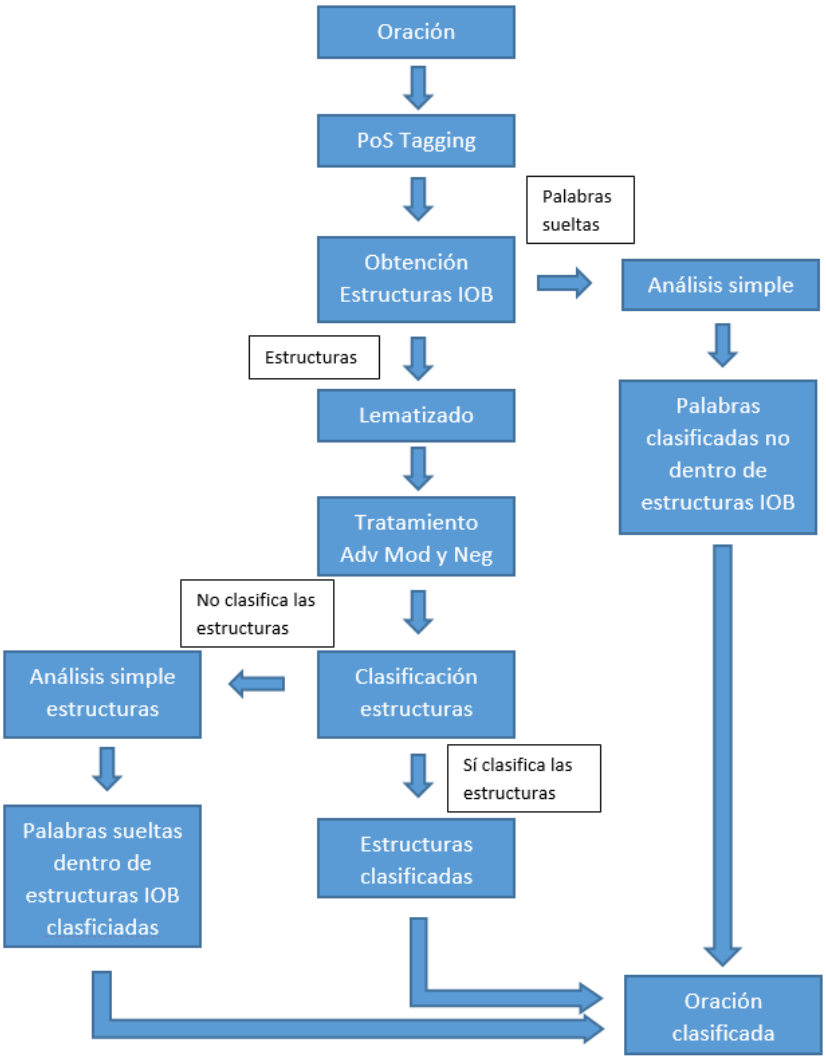



**Ilustración 9. Diagrama de análisis de estructuras IOB**

Donde se pueden observar los pasos previamente explicados, pero de una forma más visual.

A fin de facilitar la comprensión, se ha desarrollado un ejemplo, esta vez con el lexicón EmoSenticNet, que tiene más riqueza en lo que a entradas mutipalabra respecta. Vamos a utilizar la oración “Ella tiene un gran corazón”, ya que el conjunto “gran corazón” se encuentra dentro del lexicón, con las puntuaciones que se puede observar en la ilustración 10:

```
de gran corazon: [0.0, 0.0, 1.0, 0.0, 0.0, 0.0]
```

**Ilustración 10. Valores para "de gran corazón"**

Que presenta valor 1 para alegría. A la luz de estos datos, se ha realizado la ejecución del analizador, como se puede apreciar en la ilustración 11:

```
Introduzca el texto que desea analizar:
Ella tiene un gran corazón
¿Con qué lexicón quiere utilizarlo? Elija la opción deseada:
0. NRC
1. EmoSenticNet
1
[('ella', 'PRP')]
[('tiene', 'VB')]
[('un', 'DT'), ('gran', 'JJ'), ('corazón', 'NN')]
['gran', 'corazon'] [0.0, 0.0, 1.0, 0.0, 0.0, 0.0]
Resultados obtenidos con EmoSenticNet:
Enfado: 0.0
Desagrado: 0.0
Alegría: 1.0
Tristeza: 0.0
Sorpresa: 0.0
Miedo: 0.0
```

**Ilustración 11. Resultado para "un gran corazón"**

En donde se puede ver la existencia de la estructura “un gran corazón”, captada por el PoS tagger y posteriormente clasificada con los valores del lexicón.

## 6.5. Análisis con bigramas

Se trata de otra solución que contempla el análisis de entradas de más de una palabra en el lexicón, pero, a diferencia del caso anterior, el número de elementos que puede tener cada entrada está restringido a dos.

Otra diferencia es en sí el funcionamiento de aplicar la creación de bigramas. Mientras que el análisis de estructuras gramaticales divide en grupos una oración, el de bigramas crea grupos de dos con todas las palabras que son consecutivas, de manera que todas las palabras que forman parte de una oración, a excepción de la última y de la primera, formarán parte de dos bigramas: uno con la palabra que le precede, y otro con la palabra que le sigue.

De este modo, el procedimiento seguido es muy similar al anterior, si bien modificando el método de obtención de los grupos de palabras.

Este análisis de bigramas, al igual que el anterior, no se aplica a Depeche Mood por los motivos que se han comentado con anterioridad.

En la ilustración 12 se puede observar un diagrama del mecanismo detrás de este analizador:

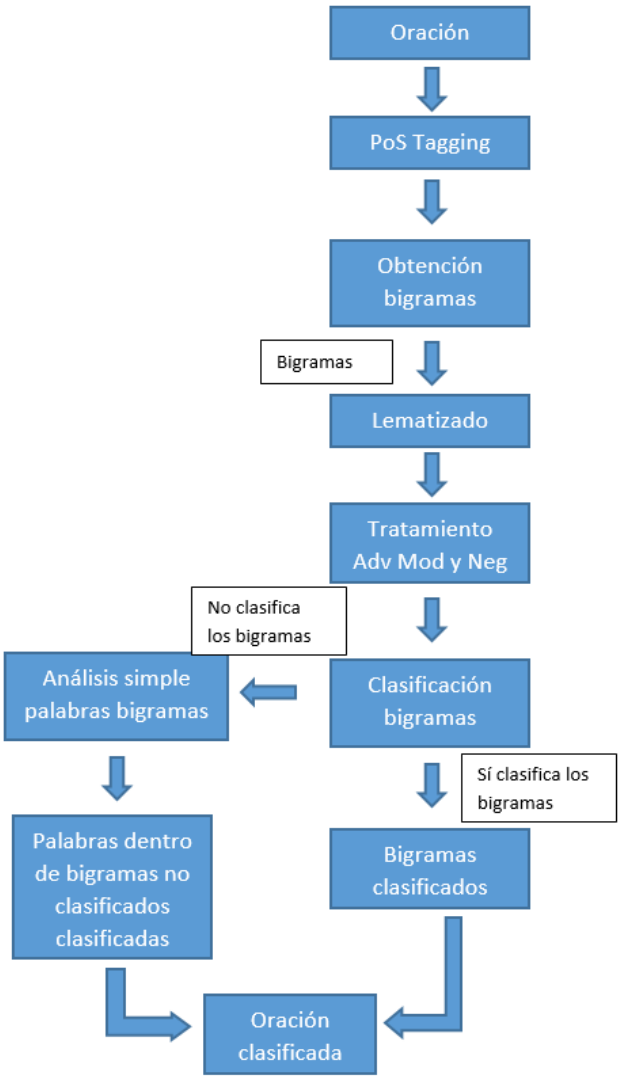



**Ilustración 12. Diagrama para análisis por bigramas.**

Por último, se va a mostrar un ejemplo que ayude a comprender mejor el funcionamiento. Utilizaremos la oración “Parece que está muy nerviosa”. Se usará el lexicón EmoSenticNet, y el valor para el conjunto “muy nerviosa” es, según lo mostrado en la ilustración 13:

```
muy nervioso: [0.0, 0.0, 0.0, 0.0, 0.0, 1.0]
```

**Ilustración 13. Valores para 'muy nervioso'**

Donde aparece un 1 en el lugar de la emoción ‘Miedo’. Teniendo esto presente, se ejecuta el programa:

```
Introduzca el texto que desea analizar:
Parece que está muy nerviosa
¿Con qué lexicón quiere utilizarlo? Elija la opción deseada:
0. NRC
1. EmoSenticNet
1
('parece', 'que')
['parece', 'que']   []
('que', 'está')
['que', 'esta'] []
('está', 'muy')
['esta', 'muy'] []
('muy', 'nerviosa')
['muy', 'nerviosa'] [0.0, 0.0, 0.0, 0.0, 0.0, 1.0]
Resultados obtenidos con EmoSenticNet:
Enfado: 0
Desagrado: 0
Alegría: 0
Tristeza: 0
Sorpresa: 1.5
Miedo: 0
```

**Ilustración 14. Resultados para análisis por bigramas**

Donde podemos ver que encuentra el bigrama “muy nerviosa” y le da su valor.

# 7. EXPERIMENTOS REALIZADOS

En este capítulo se van a explicar las diversas experimentaciones realizadas con las herramientas desarrolladas a fin de conocer su desempeño.

Para ello, se ha realizado una clasificación con cada una de las técnicas mencionadas en el capitulo 6 (Análisis simple, análisis por estructuras IOB y análisis por bigramas) usando cada uno de los lexicones (EmoLex (NRC), EmoSenticNet y Depeche Mood). Sin embargo, como para Depeche Mood no se han habilitado ni análisis por estructuras IOB ni análisis por bigramas, el número total de clasificaciones obtenidas es de 7:

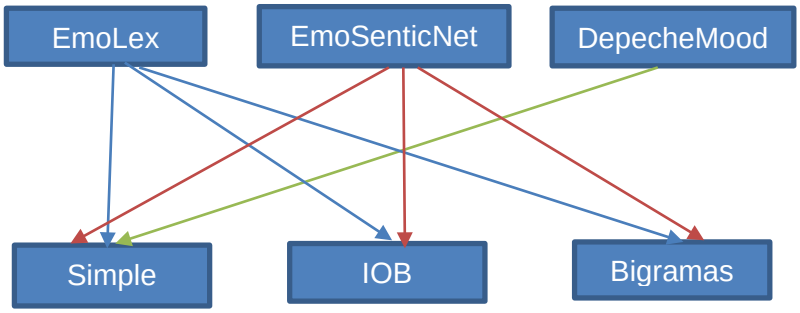



**Ilustración 15. Diagrama de combinaciones Lexicón-Estrategias**

Además, se han creado dos métodos de clasificación supervisados, uno con LinearSVC[15] y otro con Multilayer Perceptron[16], a través de la librería de Python Scikit Learn[17]. El primero se utilizará para la comparación con los resultados obtenidos para la subtarea 2 y el segundo para la subtarea 5, ambas explicadas a continuación. Para los dos métodos se ha realizado el mismo tipo de tratamiento:

1. En primer lugar, se han tratado los ficheros de entrenamiento de la siguiente manera:

[15] http://scikit-learn.org/stable/modules/generated/sklearn.svm.LinearSVC.html#sklearn.svm.LinearSVC
[16] http://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html#sklearn.neural_network.MLPClassifier
[17] http://scikit-learn.org/stable/

- Se han eliminado los emoticonos que pudiese haber presentes en el texto.
- Se han sustituido abreviaciones incorrectas comunes por sus palabras correctas. Por ejemplo, ‘q’ por ‘que’.
- Se han modificado todas las palabras relacionas con la risa con diferente longitud para que todas tengan el mismo valor. Es decir, cambiar ‘jajaja’ por ‘jaja’ y ‘xDD’ por ‘xd’.
- Se han eliminado las palabras vacías.
- Se ha lematizado el texto.
- Se han obtenido las raíces de las palabras.

2. A continuación, se han obtenido las palabras cuya ocurrencia en el conjunto de datos fuese mayor de 10 y con ellas se ha creado una bolsa de palabras. Es decir, para cada palabra de cada tweet, se coloca un cero si no está en nuestra lista de palabras con ocurrencia mayor que 10, y 1 en caso contrario.
3. Con los valores “gold” del entrenamiento y los vectores obtenidos en el apartado anterior, se entrena el modelo.
4. Tras eso, se realizan los pasos 1 y 2 con los ficheros de test.
5. Por último, se realiza la predicción.

Para evaluar los analizadores generados, se ha utilizado dos subtareas de la Tarea 1 de la competición SemEval 2018[18], concretamente las subtareas 2 y 5 [22], aquellas relacionadas con clasificación de emociones:

- Subtarea 2 (EI-oc): el objetivo de esta subtarea es el de medir la intensidad de las cuatro emociones básicas (Enfado, Miedo, Alegría y Tristeza) en un conjunto de tweets. Más adelante se profundizará en su naturaleza.
- Subtarea 5 (E-c): esta subtarea se encarga de la clasificación multietiqueta de un conjunto de tweets en 11 emociones básicas (Enfado, Anticipación, Asco, Miedo, Alegría, Amor, Optimismo, Pesimismo, Tristeza, Sorpresa, Confianza). Se profundizará en la misma a posteriori.

[18] https://competitions.codalab.org/competitions/17751

Como los resultados que devuelven los analizadores no están en esas escalas, se han tenido que normalizar los valores con diferentes técnicas para cada subtarea:

- Para la subtarea 2, se ha utilizado:
  - Percentiles (Modo 1): como se ha de realizar una escala de 4 valores (del 0 al 3), una aproximación adecuada podría ser por percentiles. Para ello simplemente se han de ordenar todas las predicciones de intensidad de una combinación y elegir las posiciones de corte que serán los percentiles.
  - Valor máximo para cada intensidad entre dos (Modo 2): la concepción inicial de esta técnica era la suma del mínimo y del máximo dividida entre 2, pero como el valor mínimo era por aplastante mayoría 0, se podía evitar el cálculo del valor mínimo. Para este método se han de facilitar los valores predictivos y los valores 'gold', y se han de almacenar los 4 valores máximos que pueden tomar los valores predictivos para los diferentes tweets que presentan cada valor 'gold' de intensidad.

    Luego se ordenan estos valores de menor a mayor para evitar incongruencias y se fijan como umbrales, teniendo 4 umbrales para cada una de las posibles intensidades. Llegados a este punto, se realiza la comparación de los valores de predicción con estos umbrales para estimar la intensidad.

- Para la subtarea 5 los métodos usados han sido los siguientes:
  - Media de los valores (Modo 1): se realiza el cálculo de la media para todos los valores de las emociones para cada tweet y se utiliza como umbral para normalizar los valores.
  - Truncamiento (Modo 2): este método es muy sencillo. Si el valor de predicción es mayor que cero, tendrá valor uno; si por el contrario vale cero, en su normalización valdrá cero.
  - Mediante cálculo de umbrales a través de la suma del máximo y el mínimo dividida entre dos o min+max/2 (Modo 3): se han de

obtener los valores mínimo y máximo para cada emoción en los valores de predicción. Luego, a partir de estos mínimos y máximos, se crean umbrales, que son la suma del mínimo y del máximo dividida entre dos:

$$Umbral = \frac{\text{Mínimo+Máximo}}{2}$$

Y se comprueba por cada valor de predicción si es mayor o igual que el umbral.

Por último, se ha de comentar que se ha producido una situación especial con la subtarea 5. Como el conjunto de emociones que se han utilizado para las comparaciones es de 5 (Enfado, Miedo, Alegría, Tristeza y Sorpresa), no se llega a los 11 de SemEval (Enfado, Anticipación, Asco, Miedo, Alegría, Amor, Optimismo, Pesimismo, Tristeza, Sorpresa, Confianza). Para ello, se han propuesto dos soluciones:

- Realizar las comparaciones con las 5 emociones: esta solución conlleva que no se puedan comparar los resultados con los publicados en la página web[19] directamente. Por este motivo se intentó contactar con algunos competidores para saber si podrían facilitar sus ficheros test y poder realizar la comparación con las 5 emociones. De todos a los que se les envió el correo, solo uno facilitó los ficheros, el grupo de la Universidad Politécnica de Valencia. A la luz de lo que podemos apreciar en la ilustración 16, se encuentra en segunda posición:

5c. E-c (multi-label emotion class.) Spanish

| # | User | Entries | Date of Last Entry | Team Name | accuracy ▲ | micro-avg F1 ▲ | macro-avg F1 ▲ |
|---|---|---|---|---|---|---|---|
| 1 | ad26kt | 4 | 01/28/18 | MILAB_SNU | 0.469 (1) | 0.558 (1) | 0.407 (2) |
| 2 | jogonba2 | 4 | 02/08/18 | ELiRF-UPV | 0.458 (2) | 0.535 (2) | 0.440 (1) |
| 3 | hhaddad | 2 | 01/08/18 | Tw-StAR | 0.438 (3) | 0.520 (3) | 0.392 (3) |

**Ilustración 16. Tres primeras posiciones del ranking de resultados de la subtarea 5 en español**

[19] https://competitions.codalab.org/competitions/17751#results

Por lo que se han podido utilizar sus resultados para la creación de un ranking interno.

Para realizar las evaluaciones, al no tener el conjunto completo de emociones, se ha tenido que realizar el cálculo de las mismas adaptando el código original[20] a nuestras necesidades.

- Crear una solución híbrida que aporte los 6 emociones restantes que le faltan a las 5 comunes para llegar a las 11 de SemEval. Así, se podría comparar con los resultados de la web, si bien esta no es la mejor opción, puesto que la acción supervisada no nos permite ver cómo funciona en su totalidad los analizadores creados.

## 7.1. Subtarea 2 (EI-oc)

### 7.1.1. Conjuntos de datos

Esta subtarea mide la intensidad en una escala de 0 a 3 (siendo 0 sin evidencia de la emoción y 3 que la emoción se ve reflejada fuertemente) de las cuatro emociones básicas (Enfado, Miedo, Alegría y Tristeza) en un conjunto de tweets en español [22].

Hay disponibles varios datasets en el sitio web[21], pero solamente usaremos dos:

- EI-oc test-gold en español: donde se encuentran los ficheros con los tweets para predicción y los valores 'gold' para cada una de las cuatro emociones. El número de elementos que posee el fichero de cada emoción es de:
  - Enfado: 628 tweets
  - Miedo: 619 tweets
  - Alegría: 731 tweets
  - Tristeza: 642 tweets
- EI-oc train en español: que contiene los ficheros para entrenamiento para cada una de las cuatro emociones. Este conjunto será utilizado para

[20] https://github.com/felipebravom/SemEval_2018_Task_1_Eval
[21] https://competitions.codalab.org/competitions/17751#learn_the_details-datasets

entrenar al clasificador supervisado que se ha creado. El número de elementos que posee cada fichero de entrenamiento es de:

- Enfado: 1167 tweets
- Miedo: 1167 tweets
- Alegría: 1059 tweets
- Tristeza: 1155 tweets

Todos los ficheros presentan la forma que podemos ver en la ilustración 17:

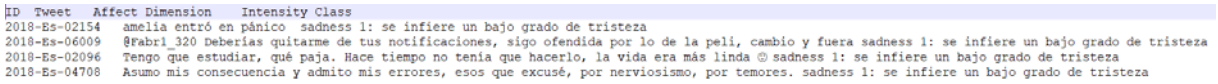



**Ilustración 17. Ejemplo de fichero El-oc para tristeza**

Donde:

- 'ID': hace referencia a la clave del tweet.
- 'Tweet': el texto de cada uno de los tweets que se van a utilizar.
- 'Affect Dimension': la emoción que se está midiendo.
- 'Intensity Class': Intensidad de dicha emoción.

### 7.1.2. Normalización

Tal y como se ha descrito en la introducción del presente capítulo, se han utilizado dos estrategias de normalización y se ha creado un clasificador supervisado para enriquecer las comparaciones.

Se ha de realizar una normalización porque la salida de los analizadores creados es el resultado de sumar las puntuaciones para esa emoción de los diferentes conjuntos (ya sean estructuras simples, estructuras IOB o bigramas) de cada tweet. Así, podemos ver un ejemplo en la ilustración 18:

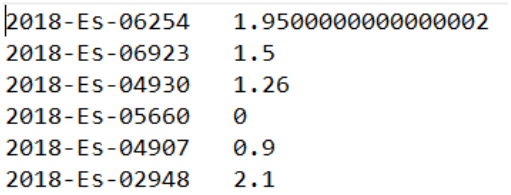



**Ilustración 18. Resultados test combinación EmoLex-Simple para Enfado**

Donde podemos apreciar que no se encuentra en la escala del 0 al 3. Por ello, para normalizar los valores en esa escala, se han utilizado las técnicas explicadas en la introducción de este capítulo, que vendrían a ser:

- Percentiles (modo 1): para los percentiles se ha seguido el procedimiento narrado en la introducción, a saber: se han ordenado de menor a mayor los valores de predicción obtenidos y se han elegido 3 percentiles como puntos de corte para lograr la normalización.

  Los percentiles escogidos han sido: 40, 52 y 75. Es cierto que los más usados son 25, 50 y 75, pero debido al elevado número de valores cero, se ha tenido que aumentar el primer percentil y ligeramente el segundo, para evitar que no valiesen dos puntos de corte lo mismo.

  Así, la asignación a la escala a través de la comparación con los umbrales se realizará de tal manera:

  - Valores menores al percentil 40: 0
  - Valores mayores o iguales al percentil 40 y menores que el percentil 52: 1
  - Valores mayores o iguales al percentil 52 pero menores que el percenil 75: 2
  - Valores mayores o iguales al percentil 75: 3
- Valor máximo para cada intensidad entre dos (Modo 2): en este método se han seguido los pasos comentados en la introducción del capítulo. Se han utilizado los valores ‘gold’ y los valores de predicción, de tal manera que se han obtenido los valores máximos para cada intensidad.

  Es decir, se ha obtenido el valor máximo de predicción para la intensidad 0 en los valores correctos, para el 1… y así con todas las intensidades.

  Una vez hecho esto, se han ordenado estos máximos de menor a mayor a fin de evitar que se pudiera producir una situación incongruente y se han dividido entre dos. Así, se han obtenido 4 umbrales para las cuatro intensidades que se normalizan de la siguiente manera:

- Valores menores que el punto 1: 0
- Valores iguales o mayores que el punto 1 pero menores que el punto 2: 1
- Valores iguales o mayores que el punto 2 pero menores que el punto 3: 2
- Valores iguales o mayores que el umbral 3: 3

### 7.1.3. Medidas de Evaluación

Las medidas de evaluación que se utilizan para esta subtarea son [22]:

- Coeficiente de correlación de Pearson: este coeficiente es una medida de relación lineal entre dos variables aleatorias cuantitativas, es decir, que se utiliza para medir el grado de relación entre estas dos variables [23].
- Coeficiente Kappa de Cohen ponderado: el Coeficiente Kappa es una medida estadística que ajusta el efecto del azar en la proporción de la concordancia observada para elementos cualitativos [24]. Concretamente, Kappa ponderado cuenta los desacuerdos (la no concordancia entre los elementos) de un modo diferente al original, asignando pesos para cuantificar la importancia relativa entre los mismos [25] [26].

De manera que se crea un ranking de resultados a través de las siguientes variantes de las dos medidas explicadas [22]:

- La medida principal es la media de los valores de correlación que se van a explicar a continuación.
- Las medidas secundarias, de las que se obtiene la media, son:
  - Coeficiente de correlación de Pearson para todas las emociones. Aparecerá como ‘Pearson 1’ en las diferentes tablas que se mostrarán en lo sucesivo.
  - Coeficiente de correlación de Pearson para un conjunto que incluye solo aquellos tweets con clases de intensidad baja, moderada y alta para cada emoción. A esta medida se le llama Coeficiente de

Correlación para algunas emociones. Equivale a ‘Pearson 2’ en las tablas posteriores.

- Coeficiente Kappa de Cohen ponderado para todas las emociones. Equivale a ‘Weighted 1’ en las siguientes tablas.
- Coeficiente Kappa de Cohen ponderado para algunas emociones. Se le llamará ‘Weighted 2’ en las tablas que se verán a continuación.

### 7.1.4. Resultados

Una vez se han dejado claras las medidas utilizadas y las diferentes metodologías para realizar la normalización, se van a mostrar los resultados obtenidos para las diferentes pruebas.

Se añadirá a las comparaciones los resultados para el método supervisado creado con LinearSVC, al que se le llamará Modo 3.

A continuación, veremos los resultados obtenidos para cada emoción en las tablas 12, 13, 14 y 15, y después se mostrará la comparación con los resultados publicados en la página de SemEval 2018[22] en la tabla 16 .

- Enfado
  - Modo 1 (Percentiles): Se han obtenido los resultados que se pueden apreciar en la ilustración 19.

[22] https://competitions.codalab.org/competitions/17751

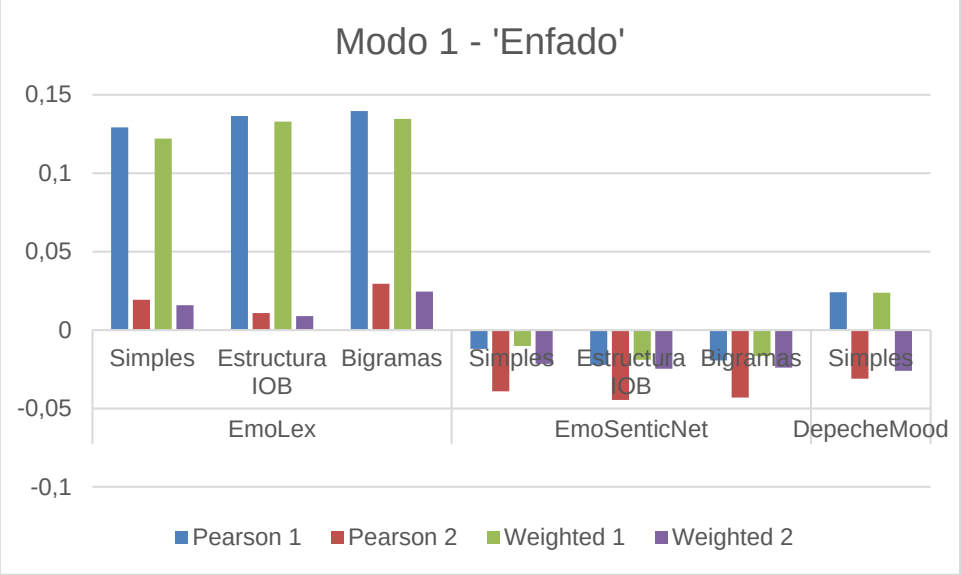



**Ilustración 19. Gráfica para el modo 1 con Enfado**

Que, analizando por cada métrica, obtenemos:

- Pearson 1: El que mejor lo ha desempeñado es el lexicón EmoLex a través de los bigramas, y el que peor EmoSenticNet a través de las estructuras IOB.
- Pearson 2: La mejor posición la gana EmoLex con bigramas y la peor es de nuevo para EmoSenticNet con estructuras IOB.
- Weighted 1: Mejor: EmoLex con bigramas, y peor EmoSenticNet con estructuras IOB.
- Weighted 2: se repiten los mismos resultados anteriores.

Como conclusión, podemos ver que, para la emoción Enfado con respecto a los percentiles, el método que mejores resultados obtiene es la conjunción del lexicón EmoLex junto con la técnica de los bigramas.

o Modo 2 (Máximos entre 2):

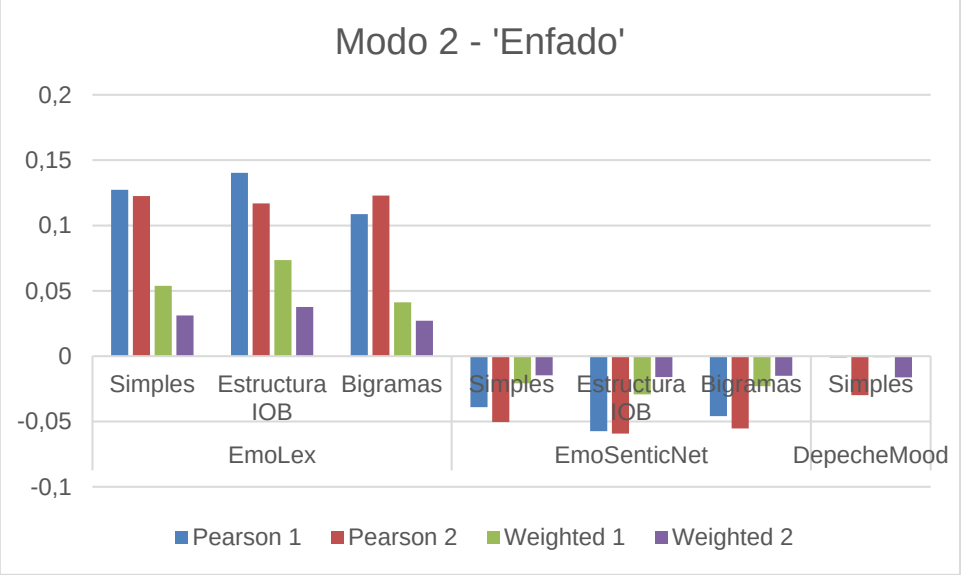



**Ilustración 20. Gráfica para el modo 2 con Enfado**

Donde podemos apreciar que:

- Para Pearson 1 el mejor es EmoLex con estructuras IOB
- Para Pearson 2 el mejor es Emolex con Bigramas
- Para Weighted 1 el mejor es EmoLex con estructuras IOB
- Para Weighted 2, también EmoLex con estructuras IOB

***Resultados para Enfado***

| *Medida* | Mejor solución | Modo 1 | Modo 2 | Modo 3 |
|---|---|---|---|---|
| *Pearson 1* | Combinación | EmoLex Bigramas | EmoLex IOB | LinearSVC |
| | Resultado | 0,139619108 | 0,140292843 | 0,32191685 |
| *Pearson 2* | Combinación | EmoLex Bigramas | EmoLex Bigramas | LinearSVC |
| | Resultado | 0,029508053 | 0,122933897 | 0,2456647 |
| *Weighted 1* | Combinación | EmoLex Bigramas | EmoLex IOB | LinearSVC |
| | Resultado | 0,134678582 | 0,073598027 | 0,30281949 |
| *Weighted 2* | Combinación | EmoLex Bigramas | EmoLex IOB | LinearSVC |
| | Resultado | 0,024575246 | 0,037729435 | 0,17924694 |

**Tabla 12. Resultados para Enfado**

A la luz de los resultados mostrados en la tabla 12, se obtiene que:

- Para todas las medidas, es el método supervisado con LinearSVC el que mejores resultados ofrece.

- En general, eliminando el supervisado, el modo 2 (máximos entre dos) es el que devuelve mejores resultados, excepto para Kappa ponderado con todas las emociones, que es ligeramente superior el modo 1 (percentiles).
- Las combinaciones con mejores resultados son EmoLex (NRC) con estructuras IOB y EmoLex con bigramas, para el modo 2.
- Para el modo 1, la mejor combinación por unanimidad es EmoLex con bigramas.

- Miedo

  - Modo 1 (Percentiles): los resultados obtenidos han sido los que se pueden observar en la ilustración 21:

Cuyos valores arrojan los siguientes resultados como mejores:

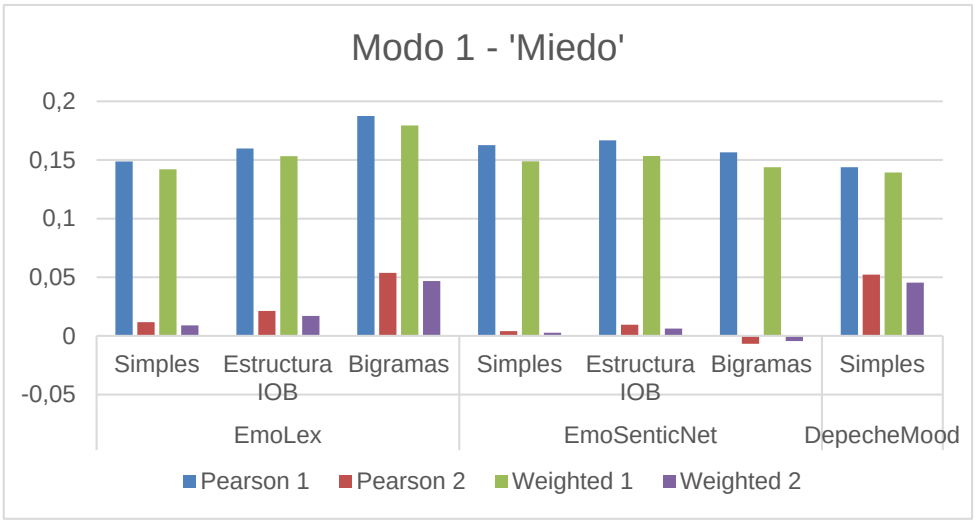



**Ilustración 21. Gráfica Modo 1 - Miedo**

- Pearson 1: EmoLex con Bigramas
- Pearson 2: EmoLex con bigramas
- Weighted 1: EmoLex con Bigramas
- Weighted 2: EmoLex con Bigramas

De donde obtenemos unánimemente que la mejor metodología para la emoción ‘Miedo’ con la modalidad de los percentiles es EmoLex con bigramas.

- Modo 2 (Máximos entre dos): según la ilustración 22:

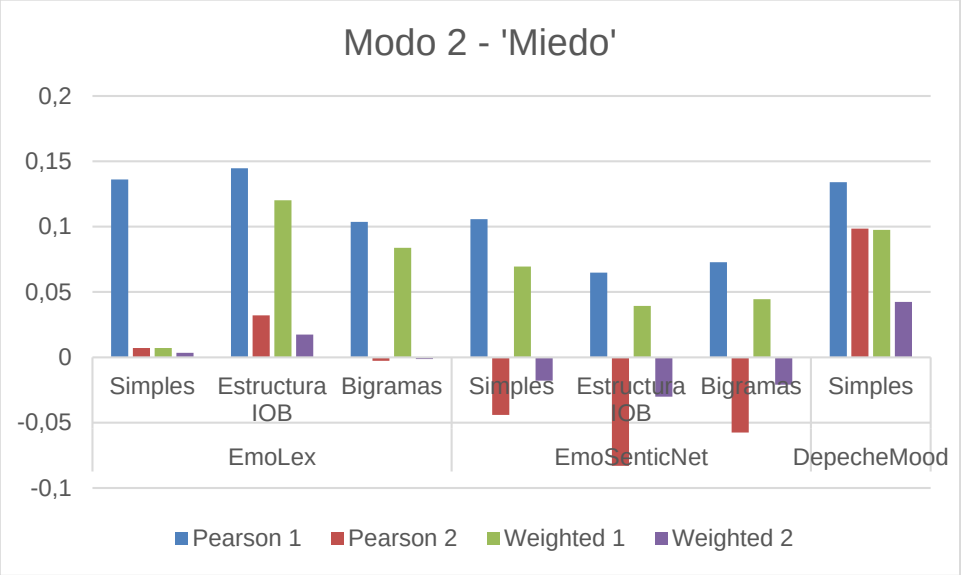



**Ilustración 22. Gráfica Modo 2 - Miedo**

- Pearson 1: EmoLex con estructuras IOB
- Pearson 2: Depeche Mood (Simple)
- Weighted 1: EmoLex con estructuras IOB
- Weighted 2: EmoLex con estructuras IOB

Viendo por primera vez aparecer como mejor opción al lexicón Depeche Mood en una de las variables, si bien por mayoría el mejor es EmoLex con estructuras IOB.

Tras analizar los resultados para cada modalidad de las relacionadas con los lexicones, se escogen las combinaciones que mejores resultados han devuelto para cada una.

***Resultados para Miedo***

| *Medida* | Mejor solución | Modo 1 | Modo 2 | Modo 3 |
|---|---|---|---|---|
| *Pearson 1* | Combinación | EmoLex Bigramas | EmoLex IOB | LinearSVC |
| | Resultado | 0,187458503 | 0,144724973 | 0,526816483 |
| *Pearson 2* | Combinación | EmoLex Bigramas | Depeche Mood | LinearSVC |
| | Resultado | 0,053698171 | 0,098529483 | 0,516833409 |
| *Weighted 1* | Combinación | EmoLex Bigramas | EmoLex IOB | LinearSVC |
| | Resultado | 0,179386018 | 0,120234517 | 0,516833409 |
| *Weighted 2* | Combinación | EmoLex Bigramas | EmoLex IOB | LinearSVC |
| | Resultado | 0,046757401 | 0,042459307 | 0,371194555 |

**Tabla 13. Resultados para Miedo**

A la luz de los resultados mostrados en la tabla 13, se obtiene que:

- Para todas las medidas, es el método supervisado con LinearSVC el que mejores resultados ofrece.
- En general, eliminando el supervisado, el modo 1 (percentiles) es el que devuelve mejores resultados, excepto para Pearson 2 (Coeficiente de correlación de Pearson para algunas emociones), que es ligeramente superior el modo 2 (media de máximos entre 2).
- Las combinaciones con mejores resultados son EmoLex (NRC) con estructuras IOB y Depeche Mood, para el modo 2.
- Para el modo 1, la mejor combinación por unanimidad es EmoLex con bigramas.

- Alegría:

  - Modo 1 (Percentiles): resultados obtenidos en ilustración 23:

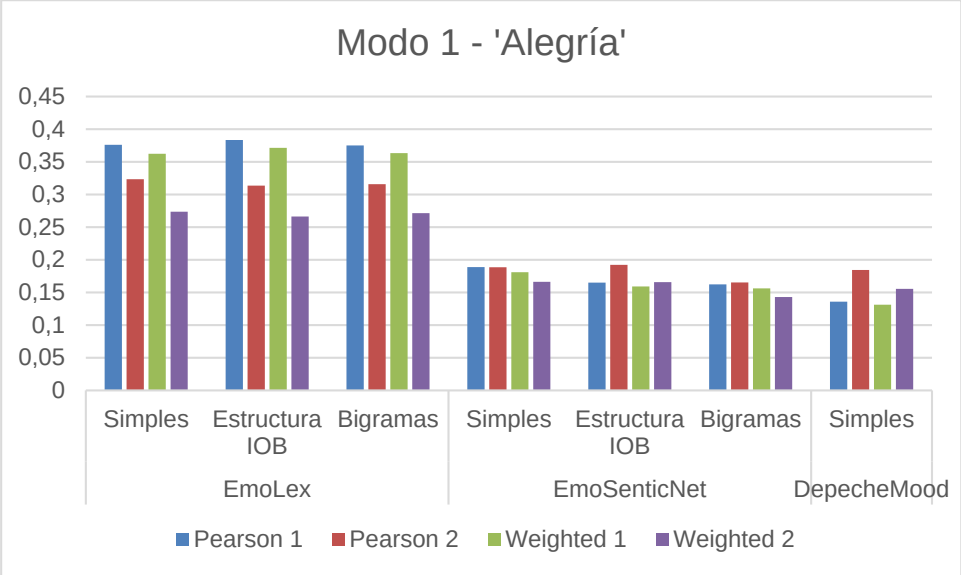



**Ilustración 23. Gráfica Modo 1 - Alegría**

- Pearson 1: EmoLex con estructuras IOB
- Pearson 2: EmoLex con estructura simple
- Weighted 1: EmoLex con estructuras IOB
- Weighted 2: EmoLex con estructura simple

En esta ocasión, se produce un "empate" entre dos técnicas diferentes, implicando ambas al lexicón EmoLex: Estructuras IOB o estructuras simples. Puesto que en el conjunto de resultados ha desempeñado mejor la opción de las estructuras simples, ésta será la opción elegida como la mejor.

- Modo 2 (Máximos entre dos): resultados obtenidos en ilustración 24.

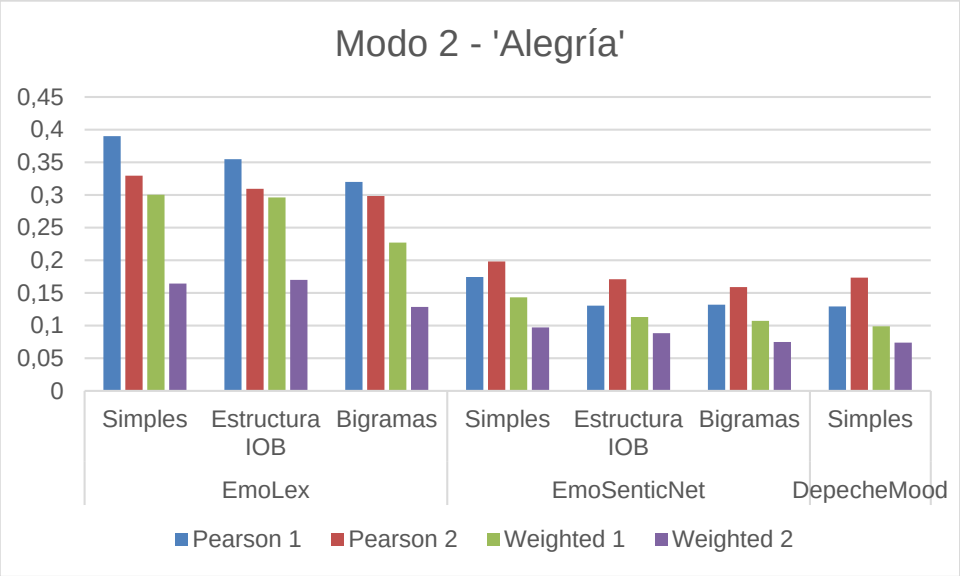



**Ilustración 24. Gráfica Modo 2 - 'Alegría'**

- Pearson 1: EmoLex con estructura simple
- Pearson 2: EmoLex con estructura simple
- Weighted 1: EmoLex con estructura simple
- Weighted 2: EmoLex con estructura simple

Donde la mejor combinación es, por unanimidad, EmoLex con estructuras simples.

Tras observar las comparaciones entre lexicones y las distintas técnicas, ahora podremos ver, a través de la tabla 14, las mejores soluciones para cada uno de los modos.

***Resultados para Alegría***

| *Medida* | Mejor solución | Modo 1 | Modo 2 | Modo 3 |
|---|---|---|---|---|
| *Pearson 1* | Combinación | EmoLex IOB | EmoLex Simple | LinearSVC |
| | Resultado | 0,383332903 | 0,390008759 | 0,48346315 |
| *Pearson 2* | Combinación | EmoLex Simple | Emolex Simple | LinearSVC |
| | Resultado | 0,323400616 | 0,32955728 | 0,37124939 |
| *Weighted 1* | Combinación | EmoLex IOB | Emolex Simple | LinearSVC |
| | Resultado | 0,371465368 | 0,300278732 | 0,47741489 |
| *Weighted 2* | Combinación | EmoLex Simple | EmoLex IOB | LinearSVC |
| | Resultado | 0,273696591 | 0,169910006 | 0,3025563 |

**Tabla 14.Resultados para Alegría**

Sobre los resultados mostrados en la tabla 14, se tiene que:

- Para todas las medidas, es el método supervisado con LinearSVC el que mejores resultados ofrece, pero en esta ocasión seguido de cerca por las otras dos soluciones.
- En general, eliminando el supervisado, el modo 1 (percentiles) es el que devuelve mejores resultados, si bien para las dos medidas de Pearson es mejor el modo 2 (máximos entre 2).
- Las combinaciones con mejores resultados son EmoLex (NRC) con estructuras IOB y EmoLex Simple, para ambas modalidades, si bien para la modalidad dos es mejor EmoLex Simple.

- Tristeza
  - Modo 1 (Percentiles): resultados observados en ilustración 25:

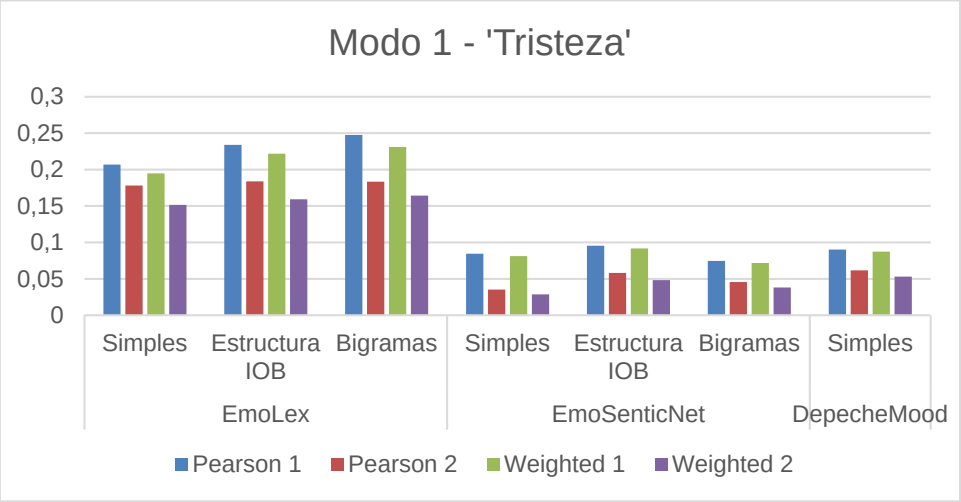



**Ilustración 25. Gráfica Modo 1 - Tristeza**

- Pearson 1: EmoLex con bigramas
- Pearson 2: EmoLex con estructura IOB
- Weighted 1: EmoLex con bigramas
- Weighted 2: EmoLex con bigramas

Obteniendo que la mejor combinación en el modo de percentiles para Tristeza es el lexicón EmoLex con bigramas.

o Modo 2 (Máximos entre dos): resultado final en ilustración 26:

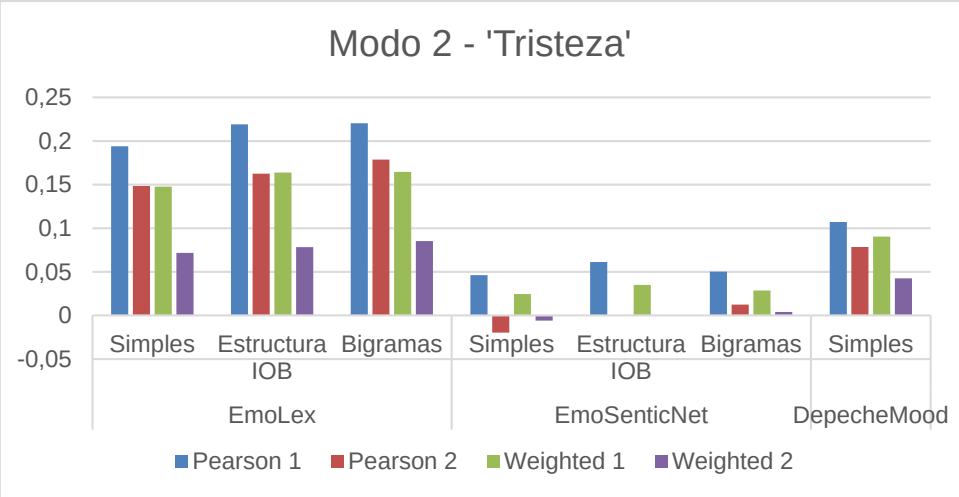



**Ilustración 26. Gráfica Modo 2 - Tristeza**

- Pearson 1: EmoLex con bigramas
- Pearson 2: EmoLex con bigramas
- Weighted 1: EmoLex con bigramas
- Weighted 2: EmoLex con bigramas

Donde por unanimidad se ve que la mejor combinación es EmoLex con bigramas.

Tras este proceso, se eligen a las mejores soluciones de ambos modos y se comparan con los resultados arrojados por el modo supervisado.

***Resultados para Tristeza***

| *Medida* | Mejor solución | Modo 1 | Modo 2 | Modo 3 |
|---|---|---|---|---|
| *Pearson 1* | Combinación | EmoLex Bigramas | EmoLex Bigramas | LinearSVC |
| | Resultado | 0,247505281 | 0,220299968 | 0,50985748 |
| *Pearson 2* | Combinación | EmoLex IOB | EmoLex Bigramas | LinearSVC |
| | Resultado | 0,183793976 | 0,178795534 | 0,4332822 |
| *Weighted 1* | Combinación | EmoLex Bigramas | EmoLex Bigramas | LinearSVC |
| | Resultado | 0,231067 | 0,16457944 | 0,49331823 |
| *Weighted 2* | Combinación | EmoLex Bigramas | EmoLex Bigramas | LinearSVC |
| | Resultado | 0,164293538 | 0,085231781 | 0,34228326 |

**Tabla 15. Resultados para Tristeza**

Según se puede observar en la tabla 15:

- Para todas las medidas, es el método supervisado con LinearSVC el que mejores resultados ofrece.
- Por unanimidad, eliminando el supervisado, el modo 1 (percentiles) es el que devuelve mejores resultados.
- Las combinaciones con mejores resultados para el modo 1 son EmoLex (NRC) con estructuras IOB y EmoLex con bigramas, si bien es mejor EmoLex con bigramas.
- Para el modo dos, es EmoLex con bigramas la mejor opción por unanimidad.

Una vez realizado el análisis por emoción, se ha de realizar un estudio comparativo con los resultados oficiales para esta subtarea, que podemos ver en la tabla 16:

***Comparación Tabla Resultados Task 2c***

| *Posición* | Usuario | macro avg 1 | macro avg 2 | macro avg 3 | macro avg 4 |
|---|---|---|---|---|---|
| *1* | Mabdou | 0,664 | 0,542 | 0,651 | 0,481 |
| *2* | IntenseEmojis | 0,599 | 0,485 | 0,581 | 0,434 |
| *6* | *Modo 3 | 0,4605 | 0,376 | 0,447 | 0,299 |
| *9* | yoyo | 0,36 | 0,331 | 0,333 | 0,27 |
| *10* | LTL_DUE | 0,342 | 0,277 | 0,253 | 0,137 |
| *11* | *Modo 1 | 0,239 | 0,147 | 0,229 | 0,127 |
| *12* | *Modo 2 | 0,224 | 0,164 | 0,164 | 0,082 |
| *17* | anongeorge | 0,007 | NaN | 0,001 | 0 |
| *18* | Random_Baseline | -0,022 | 0,016 | -0,021 | 0,015 |

**Tabla 16. Comparación resultados subtarea 2**

Donde podemos ver que, de los desarrollados, el que queda en mejor posición es el supervisado. Con respecto a los modos 1 y 2, quedan por debajo de la mitad de la tabla original, en las posiciones 11 y 12.

### 7.1.5. Análisis

Tras el estudio anterior de los diferentes resultados, y la posición final en la que quedarían las estrategias desarrolladas con lexicón, tenemos que la posición sería por debajo de la mitad de la tabla, tal y como se ha dicho, en las posiciones 11 y 12.

Lo que podemos sacar de estos resultados es que las estrategias utilizadas no han demostrado ser eficaces para medir la intensidad de las emociones, por lo que para mejorar en ese aspecto habría que sopesar otras alternativas diferentes. Otra opción sería una solución híbrida que integrara las combinaciones creadas con aprendizaje supervisado, pudiendo tener de ese modo una posible mejora en los resultados.

Otro punto a comentar es que el lexicón que por lo general mejor ha funcionado para esta subtarea ha sido EmoLex (NRC). Si miramos las técnicas con las que se puede acompañar a este lexicón, tenemos, para cada emoción:

- Enfado: con los percentiles, la mejor opción ha sido bigramas, mientras que con máximos entre dos la mejor opción ha sido estructuras IOB.
- Miedo: ambas modalidades han convergido en estructuras IOB.
- Alegría: el análisis simple ha sido el que mejores resultados ha devuelto para los dos modos.
- Tristeza: ambos modos han coincidido en la estrategia de los bigramas.

Teniendo esto claro, si se deseara utilizar alguna de las estrategias desarrolladas para la subtarea de las intensidades, las mejores opciones son las anteriormente comentadas.

## 7.2. Subtarea 5 (E-c)

### 7.2.1. Conjunto de datos

En esta subtarea se realiza la clasificación de 11 emociones básicas (Enfado, Anticipación, Asco, Miedo, Alegría, Amor, Optimismo, Pesimismo, Tristeza, Sorpresa, Confianza) sobre un conjunto de tweets en español. Los valores se presentan en binario, es decir, 0 si no presenta la emoción, 1 si sí la presenta.

Como se ha comentado al inicio del capítulo y también en el capítulo 3, en el apartado de comparación de lexicones, las comparaciones se realizan en base a las 5 emociones que tienen en común. El objetivo de esto es poder discernir cuál es mejor para la misma cantidad de emociones y en las mismas condiciones. Debido a esta situación, no se puede realizar una comparación directa con los resultados oficiales[23], por lo que se han planteado dos soluciones ya comentadas: una, adaptar la evaluación para 5 emociones, y la otra crear un híbrido que supla las emociones que faltan para comparar con los resultados oficiales. En primer lugar vamos a tratar la comparación con 5 emociones y por último la solución híbrida.

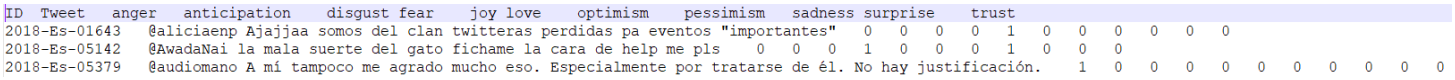
[23] https://competitions.codalab.org/competitions/17751#results

Hay disponibles varios datasets en el sitio web[24], pero solamente usaremos dos:

- E-c test-gold en español: fichero 'gold' que contiene los tweets con los que se va a realizar la parte de predicción. El número de elementos que posee el archivo es de 2855 tweets.
- E-c train en español: fichero para entrenamiento. Es utilizado para entrenar al clasificador para la solución híbrida que se ha creado. El número de elementos que posee es de 3562 tweets.

Todos los ficheros presentan la forma que podemos ver en la ilustración 27:

Donde:



| ID | Tweet | anger | anticipation | disgust | fear | joy | love | optimism | pessimism | sadness | surprise | trust |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2018-Es-01643 | @aliciaenp Ajajjaa somos del clan twitteras perdidas pa eventos "importantes" | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2018-Es-05142 | @AwadaNai la mala suerte del gato fichame la cara de help me pls | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 2018-Es-05379 | @audiomano A mí tampoco me agrado mucho eso. Especialmente por tratarse de él. No hay justificación. | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Ilustración 27. Ejemplo de fichero de subtarea 5**

- 'ID': hace referencia a la clave del tweet.
- 'Tweet': el texto de los tweets que se van a utilizar.
- Y a continuación las 11 emociones de SemEval con 0 o 1 en función de si la presenta o no.

### 7.2.2. Normalización

Tal y como se ha descrito en la introducción del presente capítulo, se han utilizado tres estrategias diferentes para abordar este problema. Estas tres metodologías se han seguido tanto para la solución híbrida como para la de 5 emociones.

Se ha de realizar una normalización porque, tal y como se ha comentado para la subtarea anterior, la salida de los analizadores creados es el resultado de sumar las puntuaciones de las 5 emociones para los diferentes conjuntos (ya sean palabras simples, estructuras IOB o bigramas) de cada tweet. Así, podemos ver un ejemplo en la ilustración 28:

[24] https://competitions.codalab.org/competitions/17751#learn_the_details-datasets

```
2018-Es-07180   [0, 0, 0, 0.9, 0]
2018-Es-00392   [0, 0, 0, 0, 0]
2018-Es-06102   [0, 0, 0, 0, 0]
2018-Es-05390   [0, 1.5, 0.9, 0, 0.9]
2018-Es-00153   [1.5, 1.5, 0, 1.5, 0]
2018-Es-00962   [0, 0, 1.1700000000000002, 0, 0]
```

**Ilustración 28. Resultados test combinación EmoLex-Simple.**

Donde se puede apreciar que no se encuentran en binario. De aquí surge la necesidad de normalizar los valores, y las técnicas que se van a emplear para ello son:

- Media (Modo 1): para cada tweet, se calcula la media de los valores de su vector de emociones. Una vez obtenida la media, se van comparando los diferentes valores del mismo vector con ella, de manera que si son menores valen 0 y si son iguales o mayores valen 1.
- Truncamiento (Modo 2): si un valor es mayor que 0, valdrá 1, y si es 0, valdrá 0.
- Cálculo de umbral a través de la suma del máximo y el mínimo dividida entre dos o minmax/2 (Modo 3): para esta modalidad se siguen una serie de pasos:
    - Se obtienen los valores máximos y mínimos para cada emoción en los valores de predicción a tratar.
    - Se crean un umbral para cada emoción, cuyo valor surge de la suma del valor mínimo y el máximo de la misma dividida entre 2.
    - Una vez tenemos estos umbrales que actuarán como puntos de corte para cada emoción, se comparan los valores de predicción con su respectivo umbral; si lo supera o iguala, su valor es 1. En caso contrario, su valor es 0.

### 7.2.3. Medidas de Evaluación

Con respecto a las medidas de evaluación que se utilizan para esta subtarea tenemos [22]:

- Exactitud multietiqueta o Índice Jaccard: debido a que se trata de una tarea de clasificación multietiqueta, se ha de utilizar una medida acorde a

ella. Se define como la división del tamaño de la intersección de los conjuntos de valores de predicción entre el tamaño de la unión [22]. Se trata de la métrica oficial para la competición, es decir, la principal.

- F1-micro: este valor es la media harmónica[25] de la precisión 'micro-averaged' y de la exhaustividad 'micro-averaged' [27]. Podemos ver su fórmula en la ilustración 29.

$$Micro-avg\ Precision\ (micro-P) = \frac{\sum_{e \in E} number\ of\ tweets\ correctly\ assigned\ to\ emotion\ class\ e}{\sum_{e \in E} number\ of\ tweets\ assigned\ to\ emotion\ class\ e}$$

$$Micro-avg\ Recall\ (micro-R) = \frac{\sum_{e \in E} number\ of\ tweets\ correctly\ assigned\ to\ emotion\ class\ e}{\sum_{e \in E} number\ of\ tweets\ in\ emotion\ class\ e}$$

$$Micro-avg\ F = \frac{2 \times micro-P \times micro-R}{micro-P + micro-R}$$

**Ilustración 29. Fórmula de F1-micro obtenida de la web de SemEval[26]**

- F1-macro: es la media de la precisión y de la exhaustividad. Su fórmula aparece en la ilustración 30.

$$Precision\ (P_e) = \frac{number\ of\ tweets\ correctly\ assigned\ to\ emotion\ class\ e}{number\ of\ tweets\ assigned\ to\ emotion\ class\ e}$$

$$Recall\ (R_e) = \frac{number\ of\ tweets\ correctly\ assigned\ to\ emotion\ class\ e}{number\ of\ tweets\ in\ emotion\ class\ e}$$

$$F_e = \frac{2 \times P_e \times R_e}{P_e + R_e}$$

$$Macro-avg\ F = \frac{1}{|E|} \sum_{e \in E} F_e$$

**Ilustración 30. Fórmula de F1-macro obtenida de la web de SemEval**

### 7.2.4. Resultados

Una vez se han dejado claras las medidas utilizadas y las diferentes metodologías para realizar la normalización, se van a mostrar los resultados obtenidos para las diferentes pruebas.

Se va a comenzar con un análisis por emociones para cada emoción de las estudiadas, a saber: Enfado, Miedo, Alegría, Tristeza y Sorpresa. Para este análisis,

[25] https://en.wikipedia.org/wiki/Harmonic_mean
[26] https://competitions.codalab.org/competitions/17751#learn_the_details-evaluation

solo se usarán dos modos, el modo 2 (truncamiento) y el modo 3(minmax/2), porque al ser una sola emoción, el método de la media no tiene sentido. Los resultados podrán verse en las tablas 17, 18, 19, 20 y 21.

A continuación se mostrarán los resultados obtenidos para la clasificación con 5 emociones usando las tres metodologías, y añadiendo las uniones de las mejores combinaciones lexicón-estrategia de análisis para cada emoción obtenidas para cada modo. Lo podremos ver en la tabla 24.

Se ha desarrollado una solución híbrida con la mejor opción de combinación (que se ha obtenido del análisis con 5 emociones que se explicará a posteriori) junto con el clasificador supervisado Multilayer Perceptron que se ha comentado al comienzo del presente capítulo. El modo de realizar el híbrido ha sido simplemente obtener las 5 emociones mediante el método con lexicón y las 6 emociones restantes con el método supervisado.

Por último, se mostrarán los resultados para la solución híbrida, en la tabla 25.

- Enfado
    - Modo 2 (Truncamiento): los resultados obtenidos los podemos apreciar en la ilustración 31.

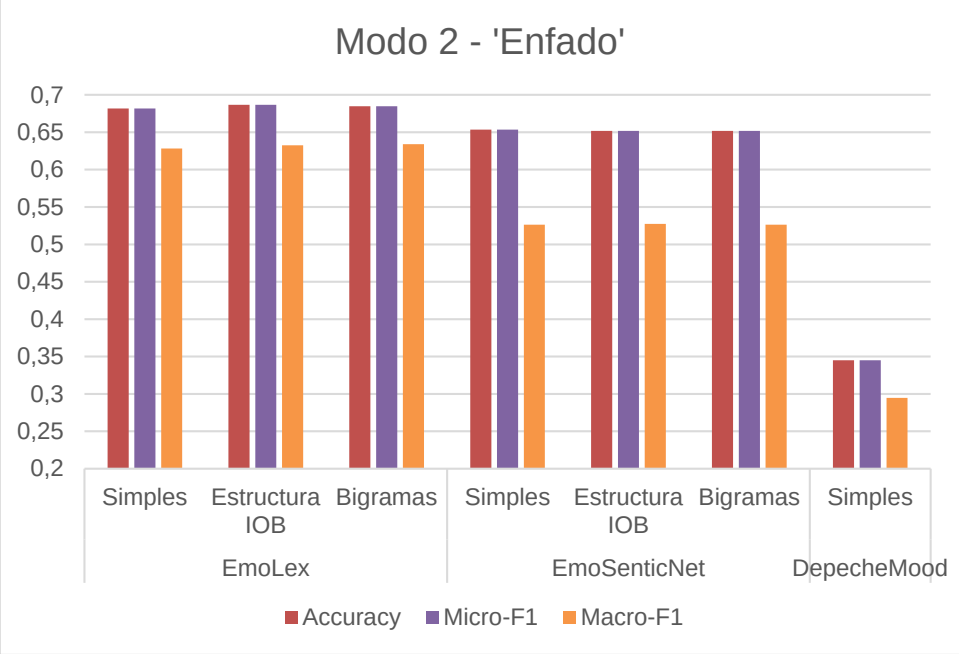



**Ilustración 31. Gráfica Modo 2 - Enfado**

- Exactitud: EmoLex con estructuras IOB
- Micro-F1: EmoLex con estructuras IOB
- Macro-F1: EmoLex con bigramas

o Modo 3 (Umbrales o Minmax): resultados observados en la ilustración 32:

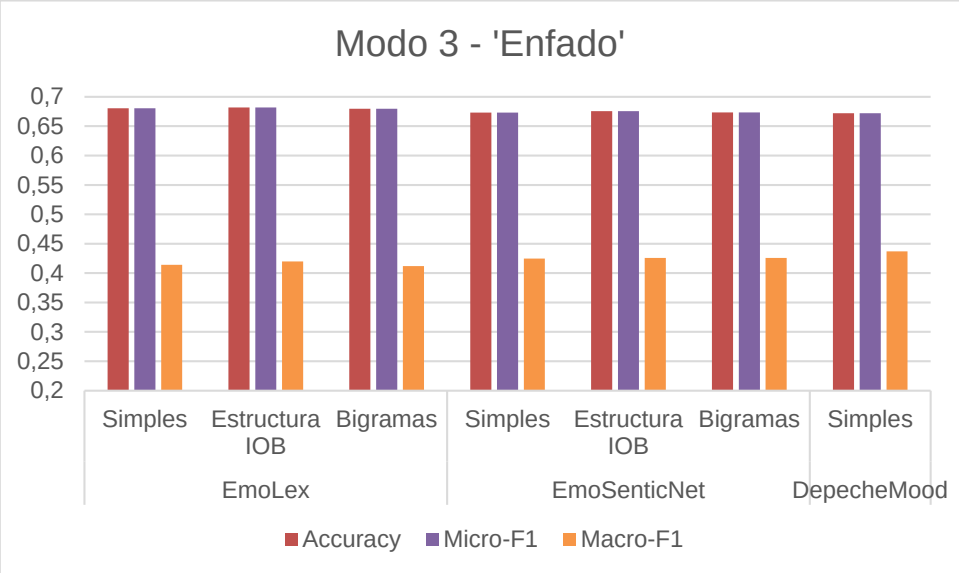



**Ilustración 32. Gráfica Modo 3 - Enfado**

- Exactitud: EmoLex con estructuras IOB
- Micro-F1: EmoLex con estructuras IOB
- Macro-F1: DepecheMood

Se escgogen los mejores resultados para cada modo y se comparan en la tabla 17.

***Resultados para Enfado***

| *Medida* | Mejor solución | Modo 2 (Trunc) | Modo 3 (Minmax/2) |
|---|---|---|---|
| *Exactitud* | Combinación | EmoLex IOB | EmoLex IOB |
| | Resultado | 0,68675543 | 0,68185004 |
| *Micro-F1* | Combinación | EmoLex IOB | EmoLex IOB |
| | Resultado | 0,68675543 | 0,68185004 |
| *Macro-F1* | Combinación | EmoLex Bigramas | DepecheMood |
| | Resultado | 0,63397075 | 0,43701681 |

**Tabla 17. Resultados para Enfado**

A la luz de los resultados mostrados en la tabla 17, se obtiene que:

- El modo 2 (Truncamiento) es el que devuelve mejores resultados.
- La mejor combinación para el Modo 2 es EmoLex (NRC) con estructuras IOB.
- Para el modo 3, las mejores combinaciones son EmoLex con estructuras IOB para exactitud y Micro-F1 y Depeche Mood para Macro-F1.

- Miedo
  - Modo 2 (Truncamiento): los resultados obtenidos los podemos apreciar a través de la ilustración 33.

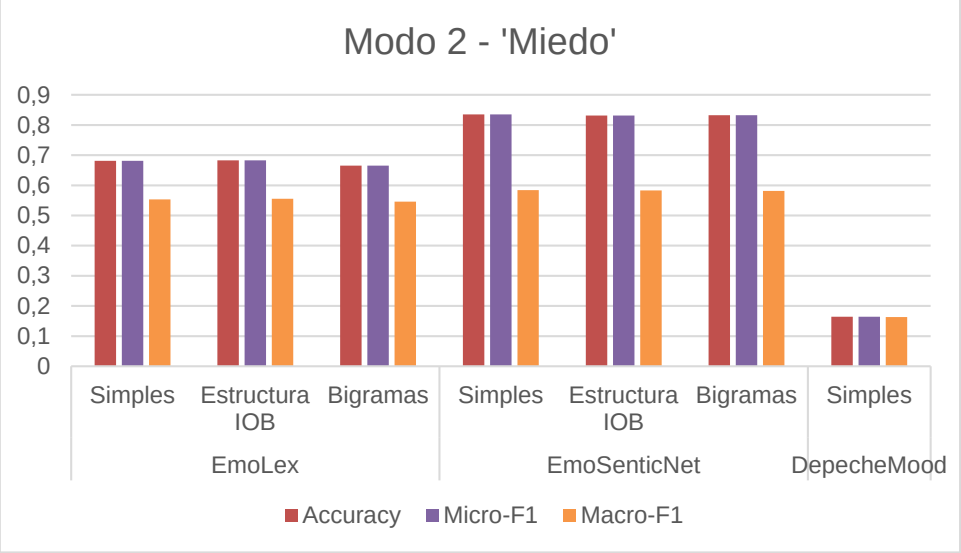



**Ilustración 33. Gráfica Modo 2 - Miedo**

- Exactitud: EmoSenticNet con análisis simple
- Micro-F1: EmoSenticNet con análisis simple
- Macro-F1: EmoSenticNet con análisis simple

o Modo 3 (Umbrales o Minmax): resultados observados a través de la ilustración 34:

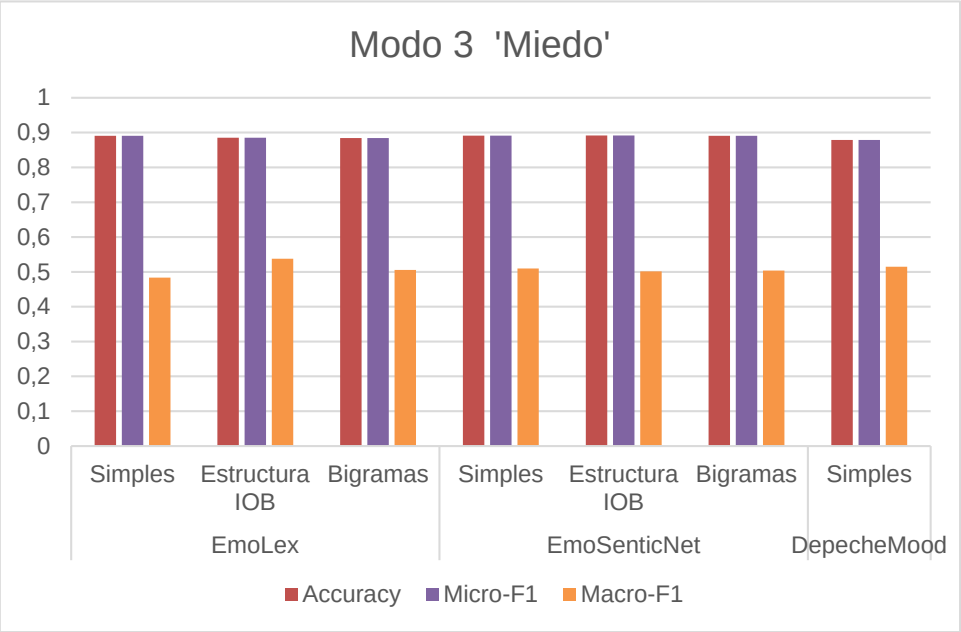



**Ilustración 34. Gráfica Modo 3 - Miedo**

- Exactitud: EmoSenticNet con análisis simple
- Micro-F1: EmoSenticNet con análisis simple
- Macro-F1: EmoLex con estructuras IOB

Tras estos resultados, comparamos los valores anteriores para cada una de las métricas:

***Resultados para Miedo***

| *Medida* | Mejor solución | Modo 2 (Trunc) | Modo 3 (Minmax/2) |
|---|---|---|---|
| *Exactitud* | Combinación | EmoSenticNet Simple | EmoSenticNet Simple |
| | Resultado | 0,83531885 | 0,89138052 |
| *Micro-F1* | Combinación | EmoSenticNet Simple | EmoSenticNet Simple |
| | Resultado | 0,83531885 | 0,89138052 |
| *Macro-F1* | Combinación | EmoSenticNet Simple | EmoLex IOB |
| | Resultado | 0,58417034 | 0,5379204 |

**Tabla 18. Resultados para Miedo**

Cuyos resultados, mostrados en la tabla 18, nos permiten apreciar que:

- El modo 3 (Minmax/2) es el que devuelve mejores resultados.
- La mejor combinación para el Modo 2 es EmoSenticNet con análisis simple.
- Para el modo 3, las mejores combinaciones son EmoSenticNet con análisis simple para exactitud y Micro-F1 y EmoLex con estructuras IOB para Macro-F1.

- Alegría

  - Modo 2 (Truncamiento): los resultados obtenidos los podemos apreciar a través de la ilustración 35.

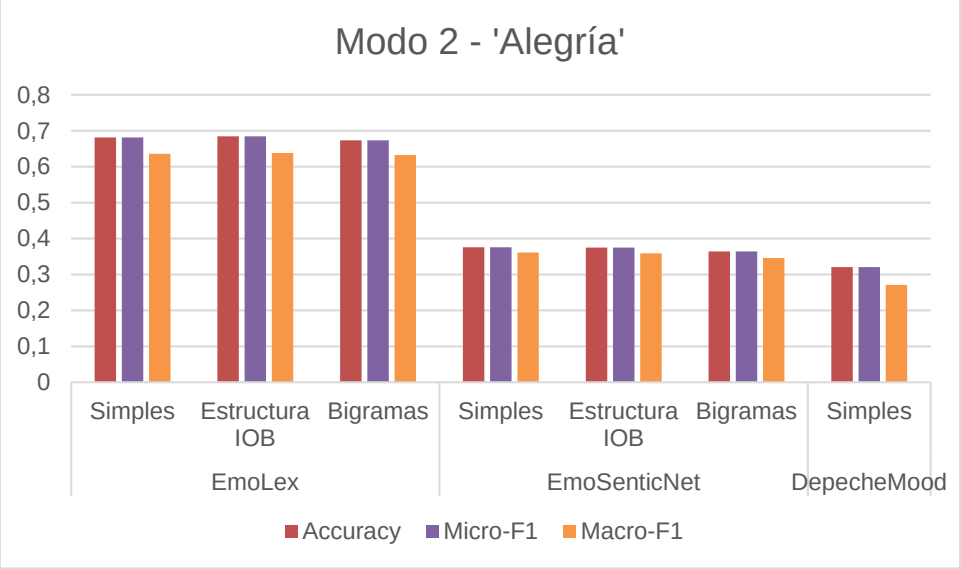



**Ilustración 35. Gráfica Modo 2 - Alegría**

- Exactitud: EmoLex con estructuras IOB
- Micro-F1: EmoLex con estructuras IOB
- Macro-F1: EmoLex con estructuras IOB

o Modo 3 (Umbrales o Minmax): resultados observados en la ilustración 36:

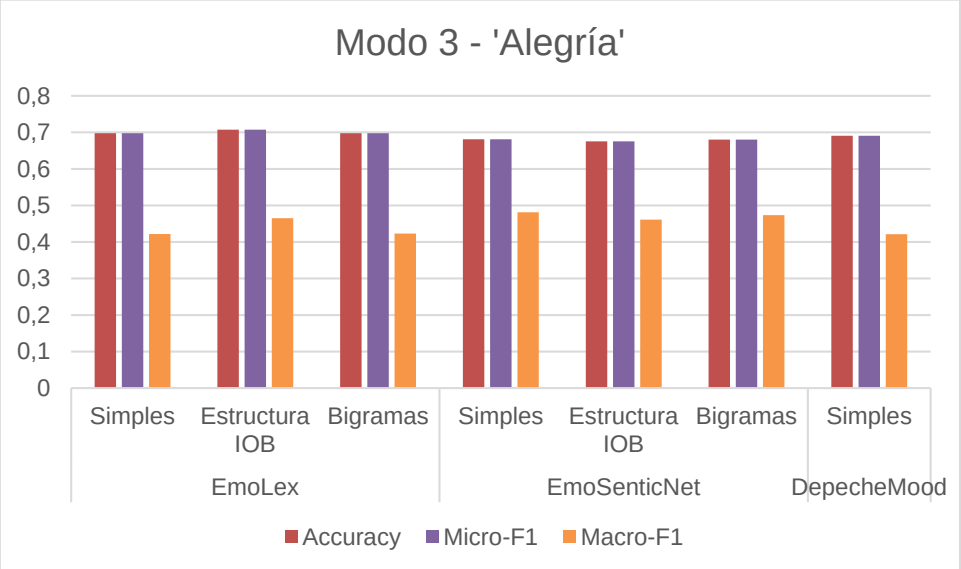



**Ilustración 36. Gráfica Modo 3 - Alegría**

- Exactitud: EmoLex con estructuras IOB
- Micro-F1: EmoLex con estructuras IOB
- Macro-F1: EmoSenticNet con análisis simple

Tras estos resultados, comparamos los valores anteriores para cada una de las métricas:

**Resultados para Alegría**

| **Medida** | Mejor solución | Modo 2 (Trunc) | Modo 3 (Minmax/2) |
|---|---|---|---|
| **Exactitud** | Combinación | EmoLex IOB | EmoLex IOB |
| | Resultado | 0,68430273 | 0,70742817 |
| **Micro-F1** | Combinación | EmoLex IOB | EmoLex IOB |
| | Resultado | 0,68430273 | 0,70742817 |
| **Macro-F1** | Combinación | EmoLex IOB | EmoSenticNet Simple |
| | Resultado | 0,63796998 | 0,48112533 |

**Tabla 19. Resultados para alegría**

Según los datos presentes en la tabla 19, tenemos que:

- El modo 3 (Minmax/2) es el que devuelve mejores resultados.
- La mejor combinación para el Modo 2 es EmoLex con estructuras IOB.
- Para el modo 3, las mejores combinaciones son EmoSenticNet con análisis simple para Macro-F1 y EmoLex con estructuras IOB para Exactitud y Micro-F1.

- Tristeza
  - Modo 2 (Truncamiento): los resultados obtenidos los podemos apreciar a través de la ilustración 37.

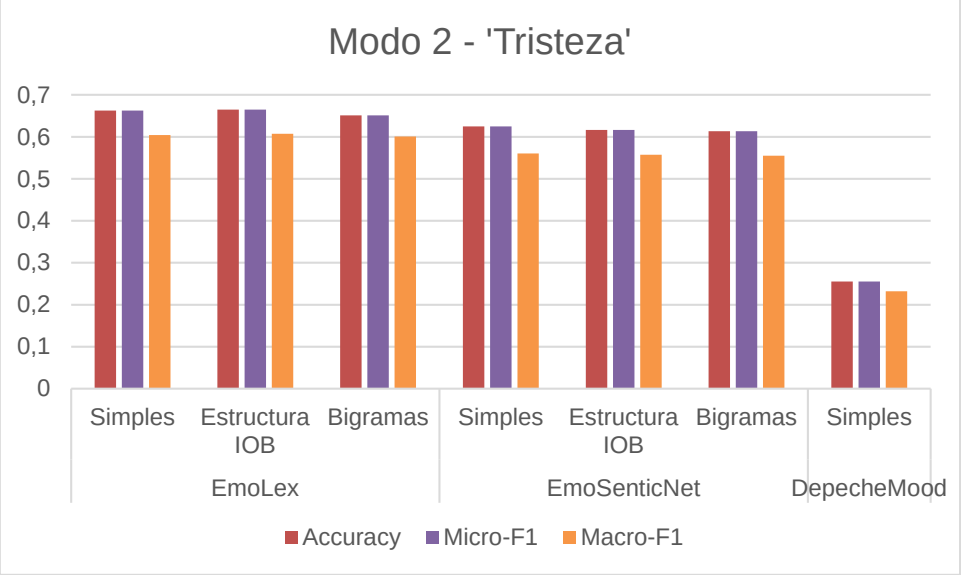



**Ilustración 37. Gráficas Modo 2 - Tristeza**

- Exactitud: EmoLex con estructuras IOB
- Micro-F1: EmoLex con estructuras IOB
- Macro-F1: EmoLex con estructuras IOB

o Modo 3 (Umbrales o Minmax): resultados observados en la ilustración 38:

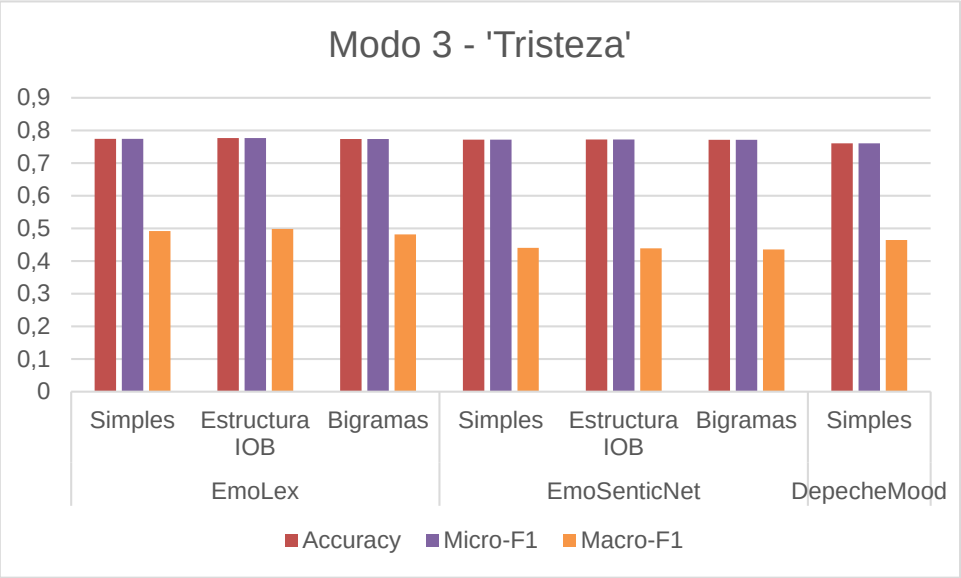



**Ilustración 38. Gráficas Modo 3 - Tristeza**

- Exactitud: EmoLex con estructuras IOB

- Micro-F1: EmoLex con estructuras IOB
- Macro-F1: EmoLex con estructuras IOB

Tras estos resultados, comparamos los valores anteriores para cada una de las métricas:

**Resultados para Tristeza**

| **Medida** | Mejor solución | Modo 2 (Trunc) | Modo 3 (Minmax/2) |
|---|---|---|---|
| **Exactitud** | Combinación | EmoLex IOB | EmoLex IOB |
| | Resultado | 0,66468115 | 0,77680448 |
| **Micro-F1** | Combinación | EmoLex IOB | EmoLex IOB |
| | Resultado | 0,66468115 | 0,77680448 |
| **Macro-F1** | Combinación | EmoLex IOB | EmoLex IOB |
| | Resultado | 0,60729486 | 0,49795523 |

**Tabla 20. Resultados para Tristeza**

A la luz de los resultados de la tabla 20, tenemos que:

- El modo 3 (Minmax/2) es el que devuelve mejores resultados, excepto en Macro-F1.
- La mejor combinación para el Modo 2 es EmoLex con estructuras IOB.
- Para el modo 3, las mejor combinación es EmoLex con estructuras IOB.

- Sorpresa

  - Modo 2 (Truncamiento): resultados obtenidos para la ilustración 39:

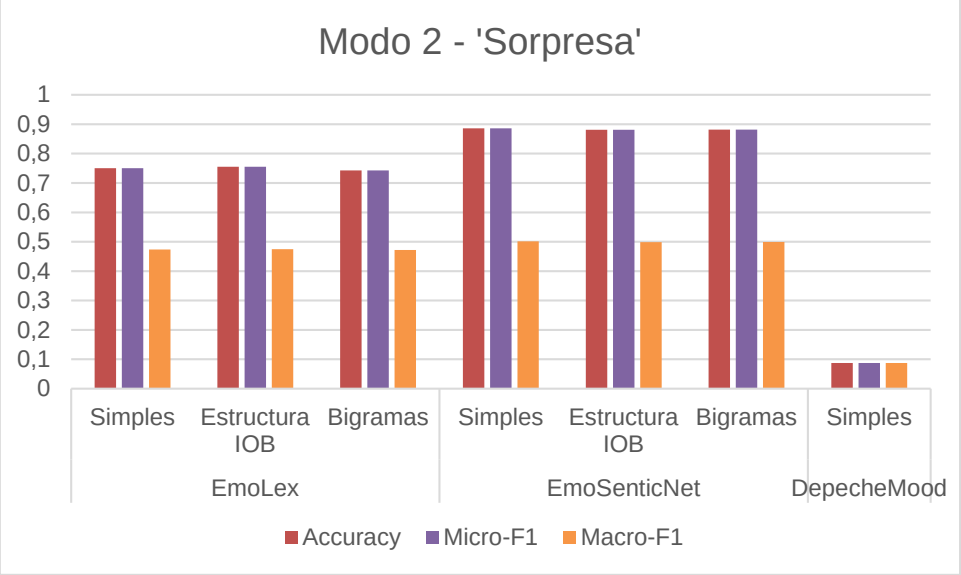



**Ilustración 39. Gráficas Modo 2 - Sorpresa**

- Exactitud: EmoSenticNet con análisis simple
- Micro-F1: EmoSenticNet con análisis simple
- Macro-F1: EmoSenticNet con análisis simple

o Modo 3 (Umbrales o Minmax): resultados observados en la ilustración 40:

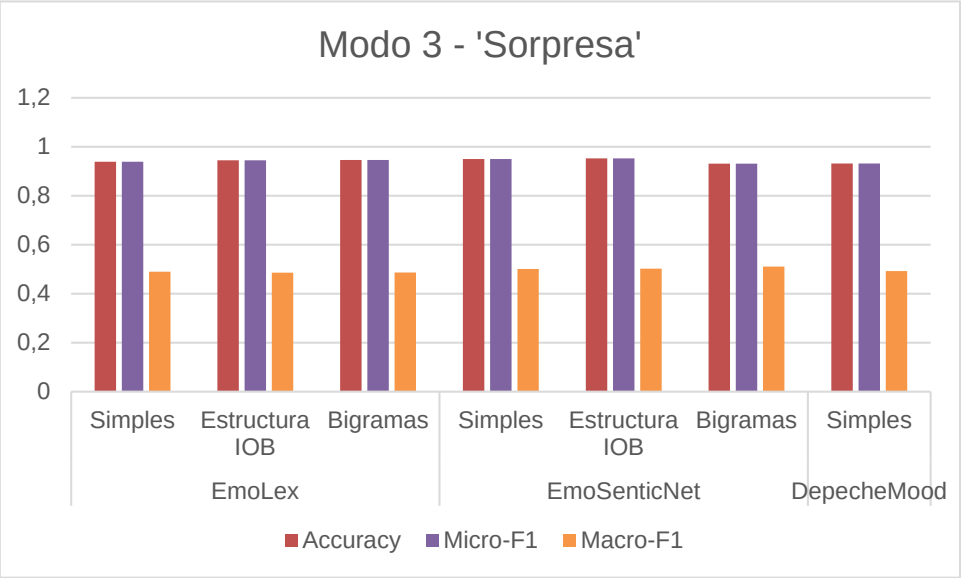



**Ilustración 40. Gráficas Modo 3 - Sorpresa**

- Exactitud: EmoSenticNet con estructuras IOB

- Micro-F1: EmoSenticNet con estructuras IOB
- Macro-F1: EmoSenticNet con bigramas

**Resultados para Sorpresa**

| **Medida** | Mejor solución | Modo 2 (Trunc) | Modo 3 (Minmax/2) |
|---|---|---|---|
| **Exactitud** | Combinación | EmoSenticNet Simple | EmoSenticNet IOB |
| | Resultado | 0,88612474 | 0,95269797 |
| **Micro-F1** | Combinación | EmoSenticNet Simple | EmoSenticNet IOB |
| | Resultado | 0,88612474 | 0,95269797 |
| **Macro-F1** | Combinación | EmoSenticNet Simple | EmoSenticNet Bigramas |
| | Resultado | 0,50138878 | 0,51056488 |

**Tabla 21. Resultados para Sorpresa**

A la luz de los resultados de la tabla 21, tenemos que:

- El modo 3 (Minmax/2) es el que devuelve mejores resultados.
- La mejor combinación para el Modo 2 es EmoSenticNet con análisis Simple.
- Para el modo 3, las mejores combinaciones son EmoSenticNet con estructuras IOB para Exactitud y Micro-F1, y EmoSenticNet con bigramas para Macro-F1.

Una vez realizado el análisis por emoción, vamos a ver cómo quedarían en un ranking las diferentes soluciones implementadas. Como el grupo de la Universidad Politécnica de Valencia facilitó sus datos, se han podido incluir en este ranking a fin de enriquecer las comparaciones.

A raíz del análisis anterior, se ha realizado la unión de las mejores combinaciones lexicón-estrategia de análisis y se ha realizado la clasificación con las mismas, con el modo 2 y el modo 3.

Podemos ver cómo está formado el modo 2 en la tabla 22:

| Enfado | Miedo | Alegría | Tristeza | Sorpresa |
|---|---|---|---|---|
| EmoLex IOB | EmoSenticNet Simple | EmoLex IOB | EmoLex IOB | EmoSenticNet Simple |

**Tabla 22. Combinaciones escogidas para modo 2 (Truncamiento)**

Y también para el modo 3 en la tabla 23:

| Enfado | Miedo | Alegría | Tristeza | Sorpresa |
|---|---|---|---|---|
| EmoLex IOB | EmoSenticNet Simple | EmoLex IOB | EmoLex IOB | EmoSenticNet IOB |

**Tabla 23. Combinaciones escogidas para modo 3 (Min Max entre dos)**

Una vez se ha aclarado la formación de ambas uniones, serán comparadas con el resto cuando se produzca la comparación de resultados en el ranking.

Con respecto a la comparaciones con las 5 emociones como conjunto para cada una de las combinaciones tenemos, tal y como se ha comentado en la introducción, tres modalidades: media (modo 1), truncamiento (modo 2) y mínimos más máximos divididos entre dos (modo 3).

- Modo 1 (Media):

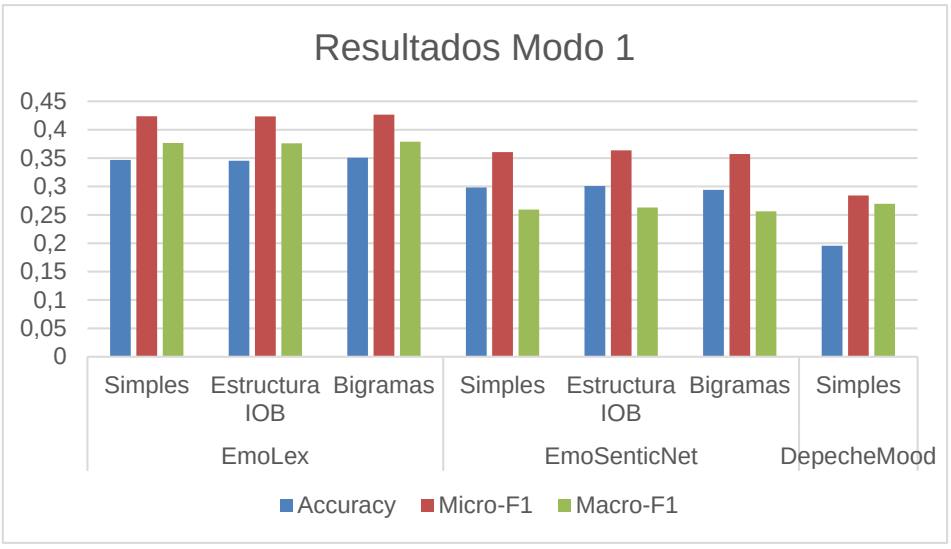



**Ilustración 41. Gráfica Modo 1 (Media)**

En la ilustración 41, podemos apreciar que el lexicón que devuelve los resultados más altos es EmoLex, en cualquiera de las tres estrategias. Le sigue el lexicón EmoSenticNet, dejando en última posición al DepecheMood.

De aquí podemos sacar que la mejor estrategia para este método de la media es EmoLex con bigramas.

- Modo 2(Truncamiento):

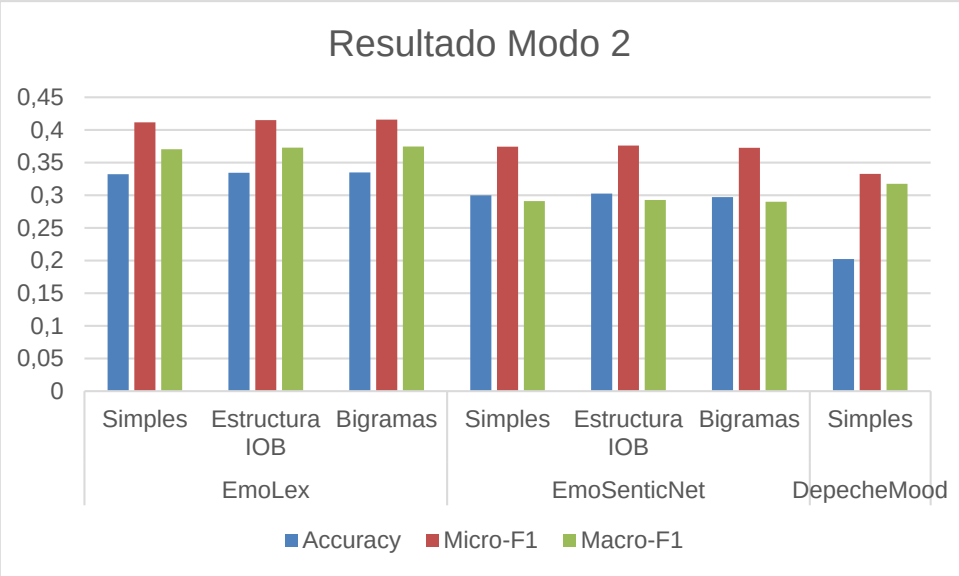



**Ilustración 42. Gráfica Modo 2 (Truncamiento)**

En la ilustración 42 se puede observar que es de nuevo EmoLex con sus combinaciones el que arroja los mejores resultados. En este caso, la mejor combinación se produce con estructuras IOB.

- Modo 3 (Umbrales o minmax):

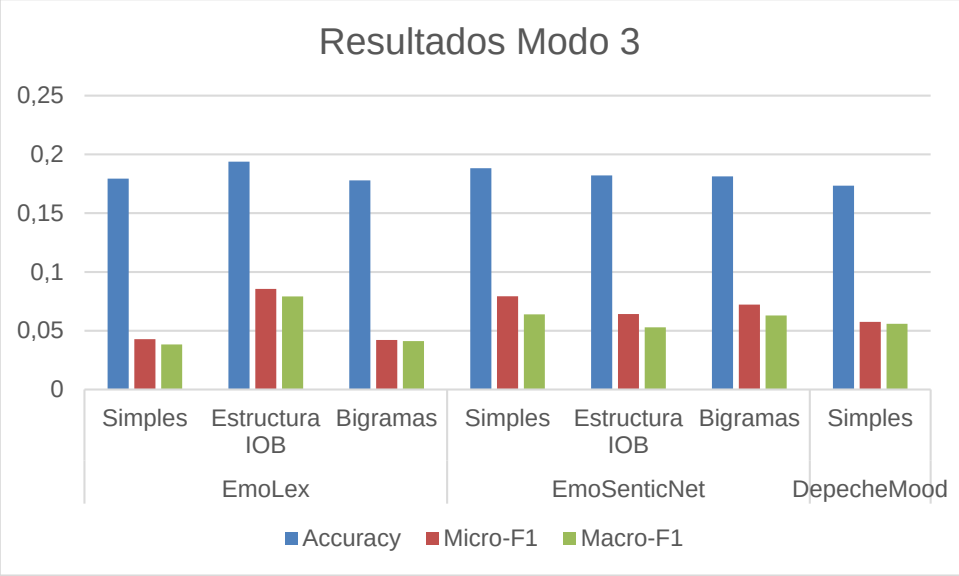



**Ilustración 43. Gráfica Modo 3 (Minmax)**

Este es el modo, según vemos en la ilustración 43, que devuelve los valores más bajos y a la vez también los más uniformes. Si bien el más alto es EmoLex con estructuras IOB, los resultados son más ajustados que en los casos anteriores.

Una vez tenemos las mejores opciones para cada modo fijo, podemos ver el resultado final en la tabla 24:

***Resultados***

| Posición | Nombre | Exactitud | Micro-F1 | Macro-F1 |
|---|---|---|---|---|
| 1 | jogonba2 | 0,57723079 | 0,64585073 | 0,55589988 |
| 2 | M2_Emo | 0,36736744 | 0,44324501 | 0,35791479 |
| 3 | M3_Emo | 0,36148096 | 0,43923174 | 0,35577574 |
| 4 | EmoLex Bigramas/M1 | 0,35054894 | 0,42666667 | 0,37890896 |
| 5 | EmoLex Bigramas/M2 | 0,33502686 | 0,41597725 | 0,3747141 |
| 6 | EmoLex IOB/M3 | 0,19390913 | 0,08554102 | 0,07919033 |

**Tabla 24. Tabla de resultados para la subtarea 5**

Donde M2_Emo es la unión de combinaciones lexicón-estrategia de análisis de la clasificación por emoción realizada anteriormente para el modo 2 (Truncamiento). M3_Emo es lo mismo, pero para el modo 3 (Min max dividido entre 2).

Como se ha mostrado en la ilustración 7, el grupo de la Universidad Politécnica de Valencia es el segundo en el ranking de resultados de SemEval.

Se puede observar que la diferencia entre los valores obtenidos por dicho grupo y la mejor combinación es alta, de 0.2, si bien en el primer caso se trata de una estrategia de aprendizaje supervisado que ha podido potenciar buenos resultados.

La mejor estrategia de las desarrolladas es M2_Emo, seguida de M3_Emo, lo que quiere decir que ha sido una buena idea el unir las mejores combinaciones.

Por último, se va a mostrar la tabla con los resultados oficiales introduciendo los de la solución híbrida:

***Resultados***

| *Posición* | Nombre | Accuracy | Micro-F1 | Macro-F1 |
|---|---|---|---|---|
| *1* | ad26kt | 0,469 | 0,558 | 0,407 |
| *2* | jogonba2 | 0,458 | 0,535 | 0,44 |
| *6* | yoyo | 0,291 | 0,354 | 0,178 |
| *7* | *Modo 2 emo | 0,287 | 0,382 | 0,302 |
| *8* | *Modo 3 emo | 0,283 | 0,381 | 0,301 |
| *9* | *Modo 1 | 0,280 | 0,376 | 0,312 |
| *10* | *Modo 2 | 0,274 | 0,372 | 0,309 |
| *11* | LTL_DUE | 0,167 | 0,275 | 0,187 |
| *12* | Random Baseline | 0,134 | 0,228 | 0,213 |
| *13* | *Modo 3 | 0,132 | 0,177 | 0,175 |
| *17* | NAVEEN,J,R | 0,05 | 0,001 | 0,001 |
| *18* | anongeorge | 0,05 | 0,001 | 0 |

**Tabla 25. Resultados oficiales para subtarea 5**

Donde se han utilizado para las 5 emociones con las que trabajábamos las distintas propuestas expuestas en el apartado anterior y para completar el resto de emociones se ha utilizado la solución hibrida.

Podemos apreciar que la mejor tentativa creada por unión con el modo de truncamiento alcanzaría la séptima posición, muy cerca de la sexta. En la octava tendríamos la versión con el modo de la suma del mínimo y el máximo entre dos.

### 7.2.5. Análisis

Con respecto al análisis por emociones, se puede determinar que ha resultado de gran utilidad. Gracias a él se ha alcanzado la mejor combinación en cuanto a resultado se refiere.

En lo que respecta a las estrategias, se puede apreciar que para esta prueba, los lexicones que mejores resultados ofrecen son EmoLex (NRC) y EmoSenticNet. En cuanto a las técnicas de análisis, las que han propiciado mejores puntuaciones han sido análisis con estructuras IOB y análisis simple.

Esta información puede resultar de gran interés si se quiere hacer un clasificador basado en lexicón a la hora de escoger estas fuentes.

A modo de conclusión, comentar que es notorio que utilizar una estrategia basada en lexicón únicamente no es la solución más eficaz. La mejor opción es una solución híbrida realizada de una manera eficaz y estudiada que permita potenciar estos resultados.

# 8. CONCLUSIONES Y TRABAJOS FUTUROS.

Una vez realizado el presente proyecto se ha podido comprobar el desempeño de los diferentes recursos léxicos que se han aplicado, teniendo presente diferentes estrategias para intentar mejorar los resultados obtenidos, y dentro de cada estrategia, utilizando diversas heurísticas para potenciar los resultados.

Como se podido comprobar, el objetivo principal de este proyecto era la generación de un léxico que nos permitiera realizar análisis de emociones en español, para lo que se ha creado la solución descrita en este documento a fin de poder suplir la necesidad en la medida de lo posible, a través de diferentes recursos que nos permitiera escoger cuál era mejor o en qué categorías para maximizar los resultados.

Teniendo en cuenta que de por sí el campo del análisis de emociones es un tema de estudio muy verde incluso en el idioma inglés, ya que es un campo relativamente reciente en la historia, lo es incluso más en nuestro idioma, ya que aunque existen muchas opciones de estudio de análisis de sentimientos que arrojan muy buenos resultados, pasar del estudio de la polaridad al de las emociones resulta en una escasez de alternativas con las que poder operar, y en el caso del castellano, opciones casi inexistentes.

Teniendo esto presente, el objetivo de la creación de este trabajo no ha sido solo el de la consecución y puesta en práctica de los conocimientos adquiridos a lo largo del proceso de aprendizaje que es este grado, sino también la posibilidad de poder ofrecer un enfoque diferente que pueda ser de utilidad.

A la luz de los resultados obtenidos, tenemos que es un campo al que le queda muchísimo por mejorar, al que le quedan muchas formas de estudio y de investigación en las que ahondar para poder ofrecer unos mejores resultados, pero que pese a ello, tal y como se comentó al inicio del presente documento, es un campo muy importante que conlleva muchísimas aplicaciones cuyo objetivo último es el de hacernos la vida más fácil y poder ayudarnos en una gran cantidad de tareas diarias o incluso en la identificación de comportamientos que pueden ser nocivos para nosotros mismos e incluso nuestra sociedad.

Se ha podido observar que, de los tres recursos léxicos utilizados, el que generalmente ha devuelto mejores resultados ha sido EmoLex de NRC, mientras que el que peores resultados ha devuelto ha sido DepecheMood. Es probable que esto se deba a que el tratamiento realizado realmente ha logrado sacar partido al primero, mientras que no ha sabido aprovechar las características del segundo.

Con respecto a los cambios y planes futuros, y siendo más concretos en el proyecto que nos atañe, una alternativa que podría mejorar sus resultados sería, sin duda alguna, el emplear una estrategia híbrida para el proceso de clasificación cuya forma de hibridación esté más estudiada y sea específica para ello. Así, una herramienta de Machine Learning podría aportar mucho al proyecto, como ya hemos podido ver en el correspondiente apartado, y de esa combinación seguramente se obtendría un buen rendimiento. Para ello, la mejor opción sería la de utilizar algoritmos con una alto desempeño, que nos puedan asegurar que se va a mejorar la solución ofrecida.

Otro de los argumentos a favor para esta integración de una combinación de estrategias de clasificación es que el análisis basado en lexicón, si bien puede otorgarnos una idea de cuáles son las principales emociones que presentan las palabras que existen en un texto, no es capaz de ir más allá con respecto a aprender e inferir las emociones reales de ese texto, cosa que con una estrategia de Machine Learning sí se podría lograr, como se ha podido apreciar a lo largo de las experimentaciones realizadas.

También estaría bien mejorar estos léxicos o incluso crearlos desde cero en el idioma español, puesto que en las traducciones, como se ha comentado anteriormente, se pierde muchas veces el valor de las emociones, ya que una palabra y su traducción no tienen por qué evocar las mismas emociones o con la misma intensidad, con lo cual esta alternativa también ayudaría a una mejora en los resultados.

Para finalizar, y de una manera más personal, concluir que el desarrollo de este proyecto ha supuesto todo un reto, que me ha permitido mejorar mis habilidades e investigar sobre materias en las que estoy altamente interesada y que realmente son,

según mi parecer, de una gran importancia para facilitarnos nuestro día a día tal y como se vive, en este mundo conectado, a día de hoy.

## Bibliografía

[1] «Idiomas por el total de habitantes, Wikipedia,» [En línea]. Available: https://es.wikipedia.org/wiki/Anexo:Idiomas_por_el_total_de_hablantes.

[2] «Real Academia Española,» [En línea]. Available: http://dle.rae.es/?id=EjXP0mU.

[3] R. d. Sousa, «Stanford Encyclopedia of Philosophy,» [En línea]. Available: https://plato.stanford.edu/entries/emotion/.

[4] J. M. Harley, «Measuring Emotions: A Survey of Cutting Edge Methodologies Used in Computer-Based Learning Environment Research,» de *Emotions, Technology, Design, and Learning*, Montréal, Canada, 2016, pp. 89-114.

[5] R. Cowie, N. Sussman y A. Ben-Ze'ev, «Emotion: Concepts and Definitions,» de *Emotion-Oriented Systems*, Belfast, Northern Ireland, UK, 2011, pp. 9-29.

[6] D. Jurafsky y J. H. Martin, «Lexicons for Sentiment and Affect Extraction.,» *Speech and Language Processing,* pp. 1-18, 2000.

[7] G. Coppin y D. Sander, «Theoretical Approaches to Emotion and Its Measurement,» de *Emotion Measurement*, Geneva, Switzerland, pp. 3-21.

[8] L. Cen, F. Wu, Z. L. Yu y F. Hu, «A Real-Time Speech Emotion Recognition System and its Application in Online Learning,» de *Emotions, Technology, Design, and Learning*, Changchun, China, 2016, pp. 27-46.

[9] R. W. Picard, «Affective Computing,» *MIT Press,* nº 1-16, 1995.

[10] A. Yadollahi, A. G. Shahraki y O. R. Zaiane, «Current State of Text Sentiment Analysis from Opinion to Emotion Mining,» *ACM Computing Surveys,* vol. 50, nº 2, pp. 1-33, 2017.

[11] O. Bruna, H. Avestiyan y J. Holub, «Emotion models for textual emotion classification,» *Journal of Physics: Conference Series,* vol. 772, nº 1, 2016.

[12] L. Canales y P. Martínez-Barco, «Emotion Detection from text: A Survey.,» 2014.

[13] B. Pătruţ y R.-P. Spatariu, «Implementation of Artificial Emotions and Moods in a Pedagogical Agent,» de *Emotions, Technology, Design, and Learning*, Bacau, Romania, 2016, pp. 63-86.

[14] «RAPPLER,» [En línea]. Available: https://www.rappler.com/ .

[15] S. M. Mohammad y P. D. Turney, «Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon,» *CAAGET ’10 Proceedings of the*

*NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text,* pp. 26-34, 2010.

[16] S. Poria, A. Gelbukh, A. Hussain, N. Howard, D. Das y S. Bandyopadhyay, «Enhanced senticnet with affective labels for concept-based opinion mining,» *IEEE Intelligent Systems,* vol. 28, nº 2, pp. 2-9, 2013.

[17] J. Staiano y M. Guerini, «DepecheMood: a Lexicon for Emotion Analysis from Crowd-Annotated News,» pp. 427-433, 2014.

[18] P. J. Stone, D. Dunphy, M. S. Smith y D. M. Oglivie, «The General Inquirer: A computer approach to content analysis.».*The MIT Press.*

[19] «SemEval, Wikipedia,» 2018. [En línea]. Available: https://en.wikipedia.org/wiki/SemEval.

[20] E. Kušen, G. Cascavilla, K. Figl, M. Conti y M. Strembeck, «Identifying emotions in social media: Comparison of word-emotion lexicons,» *Proceedings - 2017 5th International Conference on Future Internet of Things and Cloud Workshops,* 2017.

[21] M. Waala, H. Ahmed y K. Hoda, «Sentiment analysis algorithms and applications,» *Ain Shams Engineering Journal,* nº 5, pp. 1093-1113, 2014.

[22] S. M. Mohammad, S. Mohammad, F. Bravo-Marquez y S. Kiritchenko, «SemEval-2018 Task 1: Affect in Tweets,» *Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-2018),* pp. 1-17, 2018.

[23] «Coeficiente de correlación de Pearson, Wikipedia,» [En línea]. Available: https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Pearson.

[24] «Coeficiente Kappa de Cohen, Wikipedia,» [En línea]. Available: https://es.wikipedia.org/wiki/Coeficiente_kappa_de_Cohen.

[25] «Cohen's Kappa, Wikipedia,» [En línea]. Available: https://en.wikipedia.org/wiki/Cohen%27s_kappa#Weighted_kappa.

[26] «Índice Kappa con "pesos",» [En línea]. Available: http://www.hrc.es/bioest/errores_5.html.

[27] «Micro- and Macro-average of Precision, Recall and F-Score,» [En línea]. Available: http://rushdishams.blogspot.com/2011/08/micro-and-macro-average-of-precision.html.

# Anexo A. Manual de instalación.

Como se ha descrito anteriormente, el trabajo consta de dos partes diferenciadas, si bien ha habido herramientas comunes utilizadas para ambos pasos:

- Python: en primer lugar, se ha de realizar la instalación de este lenguaje, puesto que ha sido con el lenguaje con el que se ha realizado tanto la parte de generación del lexicón como la parte del analizador. La versión utilizada ha sido la versión 3.6.

- Entorno de programación para Python: como entorno de programación para Python se ha utilizado la herramienta PyCharm, puesto que es un entorno de muy buena calidad que permite realizar correctamente el trabajo.

- Pattern: herramienta utilizada para diversas utilidades de Procesamiento del Lenguaje Natural. En concreto, se han utilizado las características para el idioma español, y más concretamente para realizar el procesado de categoría gramatical (PoS tagger) que ha sido empleado en los diferentes analizadores y para la creación de la versión en español del lexicón Depeche Mood.

```
pip install git+https://github.com/clips/pattern@development#egg=pattern
```

Se instala desde ese comando para pip, puesto que descarga la versión más actualizada, ya que no se puede descargar directamente pattern en español. Para ejecutar ese comando, es necesario tener instalado git y anadir a la variable Path de las variables de entorno del sistema las siguientes rutas: C:\Program Files\Git\bin\git.exe y C:\Program Files\Git\cmd

- NLTK: el uso de esta herramienta, también relacionada con Procesamiento del Lenguaje Natural ha sido meramente simbólico, puesto que se ha utilizado en el proceso de traducción de los diferentes lexicones que se han traducido. El uso ha sido, tras instalar la utilidad necesaria para ello, el poder habilitar la obtención en español de los synsets de una palabra en inglés, para poder, de

esta manera, enriquecer el lexicón. También, para la obtención de los bigramas, se ha utilizado la herramienta proveída por NLTK.

```
pip install nltk
```

- os, re, sys y json como herramientas auxiliares utilizadas para acceder a diversas herramientas del sistema.

Con respecto a la generación de los recursos léxicos, las herramientas utilizadas han sido:

- googletrans: herramienta utilizada para la traducción con google en Python. Se realiza la instalación del paquete y ya podemos realizar traducciones a través de él.

```
pip install googletrans
```

- pydeepl: segunda herramienta utilizada para la traducción. Se basa en el traductor DeepL, un traductor que a través de redes neuronales es capaz de generar traducciones de gran calidad.

```
pip install pydeepl
```

- Watson Developer Cloud, para poder realizar el proceso de traducción a través del traductor de IBM Watson. La instalación de Watson es más compleja, puesto que se necesita de un usuario y de una contraseña, además de tener un límite de caracteres.

```
pip install watson-developer-cloud
```

Una vez se ha instalado, es a la hora de utilizarlo cuando hay que declarar las variables del usuario y la contraseña para poder crear el traductor.

- openpyxl: se ha utilizado para la lectura de ficheros xlsx, para poder obtener los datos que se encuentran dentro de ellos, procesarlos, y, en nuestro caso, transformarlos en un formato de texto de lectura más fácil para los diferentes programas.

```
pip install openpyxl
```

Para la creación de las diferentes herramientas de análisis de emociones:

- NumPy: para realizar algunas operaciones matemáticas.

```
pip install numpy
```

- sklearn: provee algunas herramientas para clasificación que han sido utilizadas en el desarrollo del modelo.

```
pip install scikit-learn
```

- lxml: para la lectura de ficheros xml se ha procedido a utilizar esta librería de Python, que promete ser mejor opción que otras más conocidas.

```
pip install lxml
```

# Anexo B. Manual de usuario.

Con respecto a la herramienta de análisis de emociones que se ha generado, a parte de una grande que nos permite elegir el lexicón y la modalidad, tenemos otras dos herramientas con los valores que mejor resultado han devuelto en la subtarea 5 explicadas en el capítulo 7.

Así, para poder utilizarlos, lo único que se ha de hacer es instalar previamente en Python los paquetes anteriomente mencionados y abrir la línea de comandos para poder movernos a la carpeta donde están situados los ficheros.

Una vez en la carpeta, ejecutamos en primer lugar “main.py”, que es el fichero que nos da la opción de elegir el lexicón y la estrategia. Así, obtendremos lo que se puede observar en la ilustración 44:

```
C:\Users\marim\Desktop\tfg2\2_Emotion_classification>python main.py
Introduzca el texto que desea analizar:
Hoy no estoy triste pero no estoy feliz
¿Con qué lexicón quiere utilizarlo? Elija la opción deseada:
0. NRC
1. EmoSenticNet
2. Depeche Mood
0
¿Qué tipo de análisis prefiere?
1. Análisis de unigramas.
2. Análisis de estructura IOB.
3. Análisis de bigramas.
3
Resultados obtenidos con el lexicón EmoLex:
Enfado: 0
Anticipación: 0.0
Desagrado: 0.6000000000000001
Miedo: 0
Alegría: 0.6000000000000001
Tristeza: 2.1
Sorpresa: 0.6000000000000001
Confianza: 0.0
```

**Ilustración 44. Resultados para main.py**

En primer lugar, se pedirá introducir el texto deseado. Una vez introducido, preguntará por el tipo de lexicón que desea utilizar. Por último, se pregunta por la estrategia a seguir, y una vez respondido, el programa muestra los resultados.

Ahora, en el fichero “modo2_main.py”, que es el que incluye la unión de las mejores combinaciones lexicón-estrategia para el modo del truncamiento en el análisis por emoción de la subtarea 5, se realizará tal y como se aprecia en la ilustración 45:

```
Introduzca el texto que desea analizar:
No me gusta nada ir a la playa
[('no', 'RB')]
[('me', 'PRP')]
[('gusta', 'VB')]
[('nada', 'DT')]
[('ir', 'VB')]
[('a', 'IN')]
[('la', 'DT'), ('playa', 'NN')]
['playa']   []
[0.36000000000000004, 0.0, 0, 0.0, 0.0, 0.36000000000000004, 0.36000000000000004, 0]
no
gusta   [0.0, 0, 0, 0, 0, 0]
nada    [0, 0, 0, 0.9, 0, 0]
ir  [0, 0, 1.5, 0, 0, 0]
playa   [0, 0, 0.9, 0, 0, 0]
[0.0, 0, 2.4, 0.9, 0, 0]
Resultados obtenidos para el modo 2 combinado:
Enfado: 0.36000000000000004
Miedo: 0
Alegria: 0.0
Tristeza: 0.36000000000000004
Sorpresa: 0
```

**lustración 45. Resultados para modo2_main.py**

Donde lo único que se ha de hacer es introducir el texto del que se quieren obtener las emociones, y ya se muestra el resultado con las 5 emociones con las que hemos trabajado.

Por último, para el fichero “modo3_main.py”, el procedimiento es el mismo que el del anterior, con la salvedad de que utiliza las mejores combinaciones para el modo de cálculo de umbrales a través de la suma de los máximos y los mínimos y su posterior división entre dos de la subarea 5. Para obtener las puntuaciones, simplemente se ha de seguir los mismos pasos que en el caso anterior. En la ilustración 46 se pueden observar los resultados:

```
Introduzca el texto que desea analizar:
Hoy estoy muy triste porque no he comido chocolate
[('hoy', 'RB')]
[('estoy', 'MD')]
[('muy', 'RB'), ('triste', 'JJ')]
['muy', 'triste']   []
[('porque', 'IN')]
[('no', 'RB'), ('he', 'MD'), ('comido', 'VBN')]
['no', 'comido']    []
[('chocolate', 'NN')]
hoy
muy
triste  []
no
comido  []
chocolate   [0, 0, 0.36000000000000004, 0, 0, 0]
[('hoy', 'RB')]
[('estoy', 'MD')]
[('muy', 'RB'), ('triste', 'JJ')]
['muy', 'triste']   []
[('porque', 'IN')]
[('no', 'RB'), ('he', 'MD'), ('comido', 'VBN')]
['no', 'comido']    [0.0, 0.0, 1.0, 0.0, 0.0, 0.0]
[('chocolate', 'NN')]
Resultados obtenidos para el modo 3 combinado:
Enfado: 0.0
Miedo: 0
Alegria: 0.0
Tristeza: 1.36
Sorpresa: 0
```

**Ilustración 46. Resultados para ejecución de "modo3_main.py"**

# Anexo C. Índice de Ilustraciones.

# Anexo D. Índice de Tablas.