

Universidad de Lima
Facultad de Ingeniería y Arquitectura
Carrera de Ingeniería de Sistemas



DETECCIÓN DE ESTADOS DE ÁNIMO MEDIANTE SENTIMENT ANALYSIS EN HISPANOHABLANTES

Tesis para optar el Título Profesional de Ingeniero de Sistemas

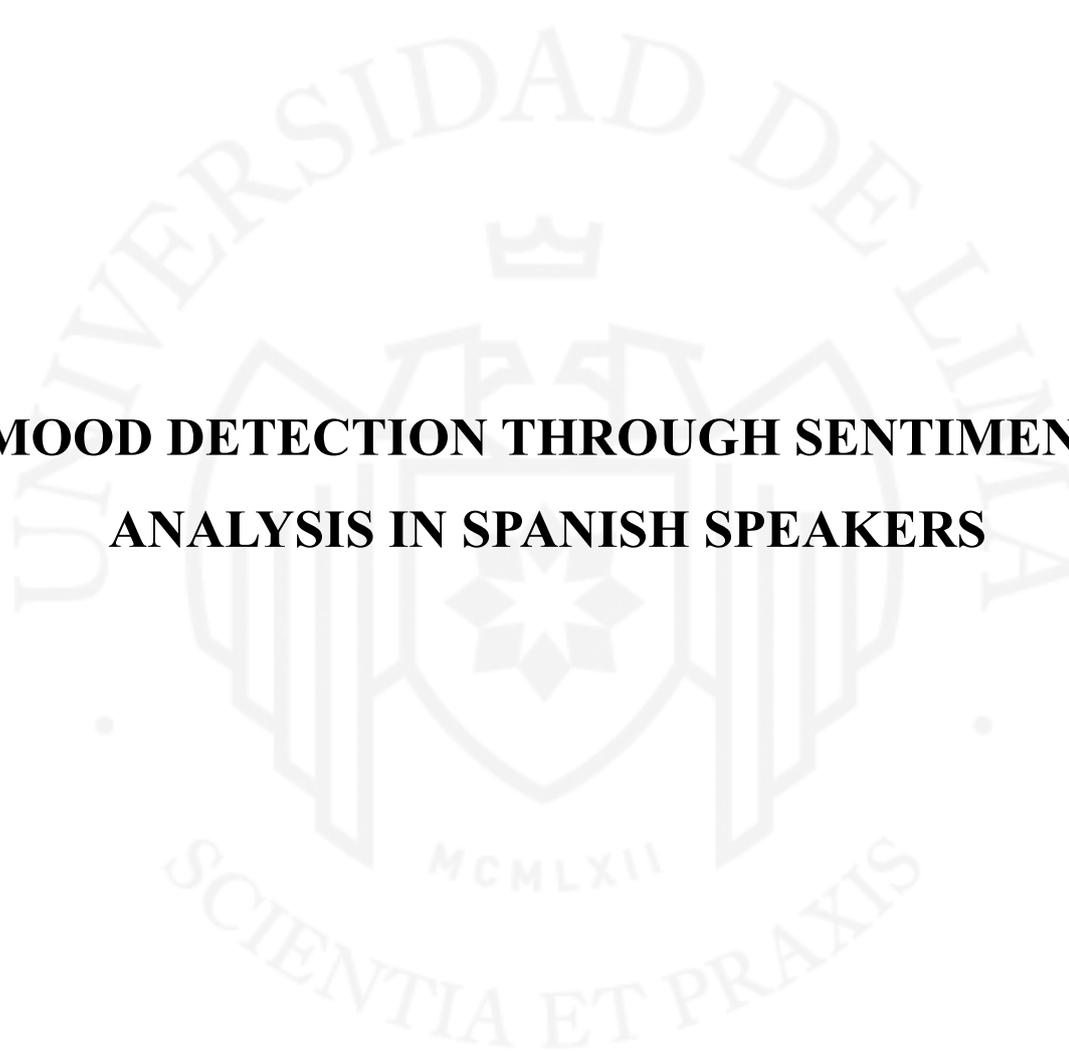
Stephany Alessandra Morzán Fuentes
Código 20120871

Asesor

Juan Manuel Gutiérrez Cárdenas

Lima – Perú
Diciembre de 2019





**MOOD DETECTION THROUGH SENTIMENT
ANALYSIS IN SPANISH SPEAKERS**

TABLA DE CONTENIDO

RESUMEN	X
ABSTRACT	XI
INTRODUCCIÓN	1
CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA	2
1.1 Formulación del problema	2
1.2 Objetivo de la investigación	4
1.2.1 Objetivo general.....	4
1.2.2 Objetivos específicos	4
1.3 Justificación	4
CAPÍTULO II: ESTADO DEL ARTE	6
CAPÍTULO III: MARCO TEÓRICO	18
3.1 Estado de ánimo	18
3.2 Emoción	23
3.2.1 Emoción y estado de ánimo	26
3.3 Natural Language Processing (NLP).....	26
3.4 Sentiment Analysis (SA).....	28
3.5 Base de datos léxica	30
3.5.1 Tesauro.....	31
3.5.2 Corpus	31
3.5.3 Lexicón	32
3.6 Keyword Spotting Technique o Técnica de Reconocimiento de Palabras Clave.....	33
CAPÍTULO IV: DESARROLLO DE LA SOLUCIÓN PROPUESTA	34
4.1 Métodos de investigación.....	34
4.2 Alcance.....	44
4.3 Supuestos.....	44
4.4 Riesgos	44
CAPÍTULO V: PRUEBAS Y RESULTADOS	45
CONCLUSIONES	54
RECOMENDACIONES	56

GLOSARIO DE TÉRMINOS	57
REFERENCIAS	58
BIBLIOGRAFÍA	64
ANEXOS	66
ANEXO 1: FORMATO DE CONSENTIMIENTO INFORMADO.....	67
ANEXO 2: URLS DEL REPOSITORIO DE CONVERSACIONES Y DEL SERVICIO IMPLEMENTADO	68
ANEXO 3: RESULTADOS DE LAS CONVERSACIONES TRATADAS POR EL SISTEMA PROPUESTO	69
ANEXO 4: VALIDACIÓN DE LA MATRIZ DE DIÁLOGO DEL BOT PARA CRITERIO DEL EXPERTO	74



ÍNDICE DE TABLAS

Tabla 2.1 Desventajas de las técnicas KST y Afinidad Léxica para categorizar textos en una emoción.....	15
Tabla 3.1 Diferencias entre emoción y estado de ánimo	26
Tabla 4.1 Frases correspondientes al saludo y pregunta inicial.....	36
Tabla 4.2 Preguntas de fondo ante respuestas positivas a la pregunta inicial	37
Tabla 4.3 Pregunta de Fondo en caso la respuesta a la pregunta inicial es negativa.....	37
Tabla 4.4 Preguntas de fondo ante respuestas indiferentes (ni positivas ni negativas) a la pregunta inicial	38
Tabla 4.5 Despedida y cierre de conversación	39
Tabla 4.6 Categorización gramatical de palabras en idioma español con Stanford CoreNLP	40
Tabla 5.1 Comparación entre los resultados de la entrevista y los resultados del Servicio Implementado	48
Tabla 5.2 Matriz de confusión de la validación de la prueba piloto	48

ÍNDICE DE FIGURAS

Figura 2.1 Relación y diferencias entre elementos subjetivos en texto.....	7
Figura 2.2 Arquitectura N° 1 para sistema de detección de emociones en texto.....	10
Figura 2.3 Arquitectura N° 2 para análisis y detección de emociones a partir de un artículo	15
Figura 2.4 Arquitectura N° 3 para detectar emociones en un texto.....	16
Figura 3.1 Visión errónea de la relación entre estados de ánimo positivos y negativos	20
Figura 3.2 Modelo bidimensional de los estados de ánimo	22
Figura 3.3 Clasificación de las emociones.....	25
Figura 3.4 Relación de Análisis de Sentimiento con NLP y con Machine Learning	29
Figura 3.5 Clasificación de Sentiment Analysis según el tipo de técnica empleada	30
Figura 4.1 Tipos de estados de ánimo reconocidos en el sistema propuesto.....	34
Figura 4.2 Arquitectura del sistema propuesto	35
Figura 4.3 Mapa de procesos del servicio del Servicio de Sentiment Analysis	40
Figura 4.4 Flujo condicional para la categorización del estado de ánimo	43
Figura 5.1 Modelo de conversación entre un participante y el bot.....	46
Figura 5.2 Estados de ánimo detectados según el sexo del participante.....	49
Figura 5.3 Proporción de la presencia de los tipos de estados de ánimo en los participantes	50
Figura 5.4 Palabras del lexicon emocional asociadas a estados de ánimo positivos más utilizadas	51
Figura 5.5 Palabras del lexicon emocional asociadas a estados de ánimo negativos más utilizadas	52

ÍNDICE DE ALGORITMOS

Algoritmo 2.1 Algoritmo para determinar el valor de una emoción en un texto.....	12
--	----



ÍNDICE DE ANEXOS

ANEXO 1: Formato de consentimiento informado.....	67
ANEXO 2: URLs del repositorio de conversaciones y del servicio implementado.....	68
ANEXO 3: Resultados de las conversaciones tratadas por el sistema propuesto	69
ANEXO 4: Validación de la Matriz de Diálogo del bot para criterio del experto	74



RESUMEN

Esta investigación muestra la elaboración y validación de un modelo computacional diseñado para detectar los estados de ánimo, definidos en la teoría de Christophe André (2010), en hispanohablantes a partir de respuestas que ingresaron estos usuarios en formato de texto. Se emplea la Técnica de Reconocimiento de Palabras Clave –en inglés, Keyword Spotting Technique (KST) – para el tratamiento del texto, y la técnica de Análisis de Sentimiento –en inglés, Sentiment Analysis– basado en solo los conceptos del Procesamiento de Lenguaje Natural –en inglés, Natural Language Processing (NLP)– para el análisis textual. Se utilizó para la recopilación de conversaciones una aplicación móvil con interfaz de chatbot, y un bot que invita al usuario a dar detalles sobre su estado de ánimo a través de preguntas validadas por un experto. Como resultado de la prueba piloto de 30 conversaciones, y luego de la prueba final en la muestra de 49 conversaciones, se obtuvo una correcta clasificación del 70% de los casos, así como una propuesta para el tratamiento de textos en idioma español.

Palabras clave: Estado de ánimo, Procesamiento de Lenguaje Natural (NLP), Sentiment Analysis, Base de datos léxica, Keyword Spotting Technique (KST).

ABSTRACT

The following research shows the elaboration and validation of a computational model designed to detect Spanish speakers' moods from answers entered by these users in text format. It was used the Keyword Spotting Technique (KST) for text treatment, and Sentiment Analysis based on only Natural Language Processing (NLP) concepts for text analysis. A mobile application with chatbot interface and a bot that invites the user to give details about their mood through questions validated by an expert, are the tools used to collect all the conversations. As a result of the pilot test of 30 conversations, and after the final test in the sample of 49 conversations, there is a correct classification of 70% of the cases, as well as a proposal for the treatment of texts in Spanish.

Keywords: Mood, Natural Language Processing (NLP), Sentiment Analysis, Lexical Database, Keyword Spotting Technique (KST)

INTRODUCCIÓN

¿Quién no ha experimentado uno de esos momentos alegres y gratificantes que da gusto contarlos?, como también ¿quién no ha pasado por alguna situación triste o incómoda que prefiere olvidar? Estas reacciones que se producen frente a determinadas situaciones son las que determinan el “estado de ánimo”, que según Thayer (1998) es el reflejo de todo aquello que le impacta a una persona.

Estas conductas pueden abordarse desde diferentes enfoques, uno de ellos es el de la ciencia de la información, la cual evoluciona con el tiempo permitiendo el desarrollo de estudios interdisciplinarios que indagan en la conducta humana, y cuyos resultados repercuten en las personas. Precisamente, el enfoque de investigación de la Inteligencia Artificial, que nació hace aprox. 60 años, ha ido cambiando conforme cambiaban las personas: sus necesidades, su interacción con su entorno y su visión del mundo (Darlington, 2017).

Es así que existen dispositivos y sistemas que logran detectar y medir el estado de ánimo de una persona, entre estos elementos se encuentran aquellos que evalúan el estado de ánimo mediante la autoexpresión, es decir, permiten que una persona exprese cómo se siente en ese momento por medio del lenguaje natural de manera textual (Desmet, 2015).

Uno de los principales desafíos de la Inteligencia Artificial con relación al procesamiento del lenguaje natural es lograr que las tecnologías de información comprenda a su locutor y responda coherentemente (Carbajales, 2018). El idioma español, como afirman Verdejo y Gómez (como se citó en Carbajales, 2018) presenta mayores complejidades que el idioma inglés, pues presenta mayores cambios según la variedad regional y, además dentro de cada variedad, existen diversas formas de uso del español.

CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA

1.1 Formulación del problema

En esta era digital existen diversas aplicaciones web y móviles para un sinnúmero de usos en las diferentes áreas y ámbitos del quehacer humano, como las aplicaciones que contribuyen con el autoconocimiento del estado de ánimo: Daylio (Beenish, 2016), Felt (Domenach, Charmarai, Savva, y Christou, 2015), Stigma (Brafman, 2017), Mood Meter App (Brackett y Stern, 2017), MyMoodTracker (Aspyre Solutions, 2010), MoodPanda (Larter, 2012), Moodtrack Social Diary (Windwer, 2017), Registro de Estados de Ánimos (AR Productions Inc., 2018), Happiness (Good to Hear, 2012), Track Your Happiness (Wong, 2015), entre otros. Precisamente, una de las ramas de la Inteligencia Artificial (IA), como lo indica Goasduff (2018), está orientada a la detección, análisis, procesamiento, interpretación y respuesta de las emociones y estados de ánimo de las personas: la IA Emocional. Para detectar los estados de ánimo, estos sistemas emplean una serie de métodos y tecnologías que permiten solicitarle al usuario el registro de los datos que estos sistemas necesitan para su funcionamiento, y que luego se procesa para finalmente generar información. Es así que existe un conjunto de aplicaciones que requieren que el usuario registre una respuesta por medio de una interfaz para proceder con su análisis.

Estas interfaces, para el registro de respuestas, pueden emplear emoticones predefinidos, como Daylio (Beenish, 2016) o Felt (Domenach et al., 2015); o pueden usar palabras previamente determinadas, como Stigma (Brafman, 2017). También se puede lograr por medio de un rango de colores, como Mood Meter App (Brackett y Stern, 2017); o mediante una calificación de la escala del 1 al 5 o del 1 al 10, más una breve reseña, como MyMoodTracker (Aspyre Solutions, 2010), MoodPanda (Larter, 2012) o la app Registro de Estados de Ánimo (AR Productions Inc., 2018). Incluso, pueden presentarse como un diario personal y registrar texto de forma ilimitada, tal es el caso de Happiness (Good to Hear, 2012; Quantified Self on Happiness, 2011). Asimismo, existen las que envían preguntas por medio de notificaciones para que el usuario las responda de manera textual, como lo hace Track Your Happiness (Wong, 2015).

Así como cada interfaz tiene su mecánica para recibir datos, también tienen su propia forma de mostrar el historial de la información analizada, e incluso la opción de compartirla. Por ejemplo, la aplicación MoodTrack Social Diary emplea un gráfico de líneas para mostrar el cambio anímico por fechas y cada hito está marcado por el color del estado de ánimo anotado (Windwer, como se citó en Google Play, 2018).

Ante esta información acerca de aplicaciones móviles que detectan estados de ánimo, surge la siguiente interrogante: ¿por qué las personas lo emplean? Según lo formulado por los usuarios de estas aplicaciones, las utilizan para conocer y determinar el comportamiento de sus propios estados de ánimo (Caldeira, Chen, Chan, Pham, Chen, y Zheng, 2017).

Sin embargo, de las aplicaciones mencionadas solo dos soportan idioma español, y todas son dirigidas a un público de idioma anglosajón, pues su primera o única versión lanzada a los mercados de aplicaciones se hizo en idioma inglés (Aspyre Solutions, 2010; Beenish, 2016; Brackett y Stern, 2017; Brafman, 2017; Domenach et al., 2015; Good to Hear, 2012; Larter, 2012; Quantified Self on Happiness, 2011; Windwer, 2017; Wong, 2015). Ninguna fue diseñada para países de habla hispana; por consiguiente, quienes solo hablan español no tienen la posibilidad de utilizar estas herramientas.

Esta dificultad hace perder la posibilidad de probar la capacidad y efectividad de estas aplicaciones en estas comunidades; una de ellas es la comunidad de hispanohablantes peruanos. En Perú, el 80% de sus habitantes hablan en español (Blasco, 2018) y está considerado, según la firma internacional de analítica y asesoría Gallup (2018), como uno de los tres países más emocionales del mundo de acuerdo al último Global Emotions Report; por ello, la implementación de estos sistemas resulta necesario y atractivo en el mundo científico.

1.2 Objetivo de la investigación

1.2.1 Objetivo general

Desarrollar un modelo para sistemas de detección de estado de ánimo cuya data de entrada sean respuestas textuales, y que esté dirigido en primera instancia a usuarios hispanohablantes.

1.2.2 Objetivos específicos

- Diseñar una estructura que soporte la detección de estados de ánimo cuya data de entrada sean respuestas textuales.
- Implementar la estructura diseñada para la construcción de un sistema de detección de estados de ánimo cuya data sean respuestas textuales y esté dirigido en primera instancia a usuarios hispanohablantes.
- Validar la estructura implementada en un sistema de detección de estados de ánimo con respuestas textuales ingresadas por usuarios hispanohablantes.

1.3 Justificación

Presentar un modelo que logre procesar respuestas textuales ingresadas en idioma español para usuarios hispanohablantes que utilicen un sistema de detección de estados de ánimo es trascendente para el estudio y tratamiento de las complejidades del lenguaje natural en idioma español, así como para el estudio de la conducta humana desde el enfoque de investigación de la ciencia de información. Es importante porque incrementa la iniciativa de desarrollar dispositivos y sistemas para hispanohablantes que buscan su autoconocimiento de forma cotidiana y desean que los datos que ingresen se registren, analicen y procesen en el idioma que se comunican.

Se contribuye así con el desarrollo de modelos y sistemas que logren efectuar métodos propios del procesamiento de la lengua española, pues los procesos y algoritmos desarrollados para otro idioma no funcionan correctamente al aplicarlo al español (Verdejo, como se citó en Carbajales, 2018). De esta manera, se aporta en el estudio de la IA Emocional al detectar emociones en las personas a través de sus

diferentes expresiones. Por lo tanto, las diversas disciplinas que estudian las actividades humanas pueden beneficiarse de los hallazgos que se obtengan.



CAPÍTULO II: ESTADO DEL ARTE

El Procesamiento del Lenguaje Natural (NLP) y Machine Learning (ML) son áreas de estudio que provienen de la IA, que se interrelacionan para brindar soluciones (Karimkhan, 2017). Mientras que ML brinda a un sistema computacional la capacidad de aprender la gestión de datos sin necesidad de intervención humana, el NLP permite que el sistema pueda interpretar el lenguaje de las personas (Michael, 2018).

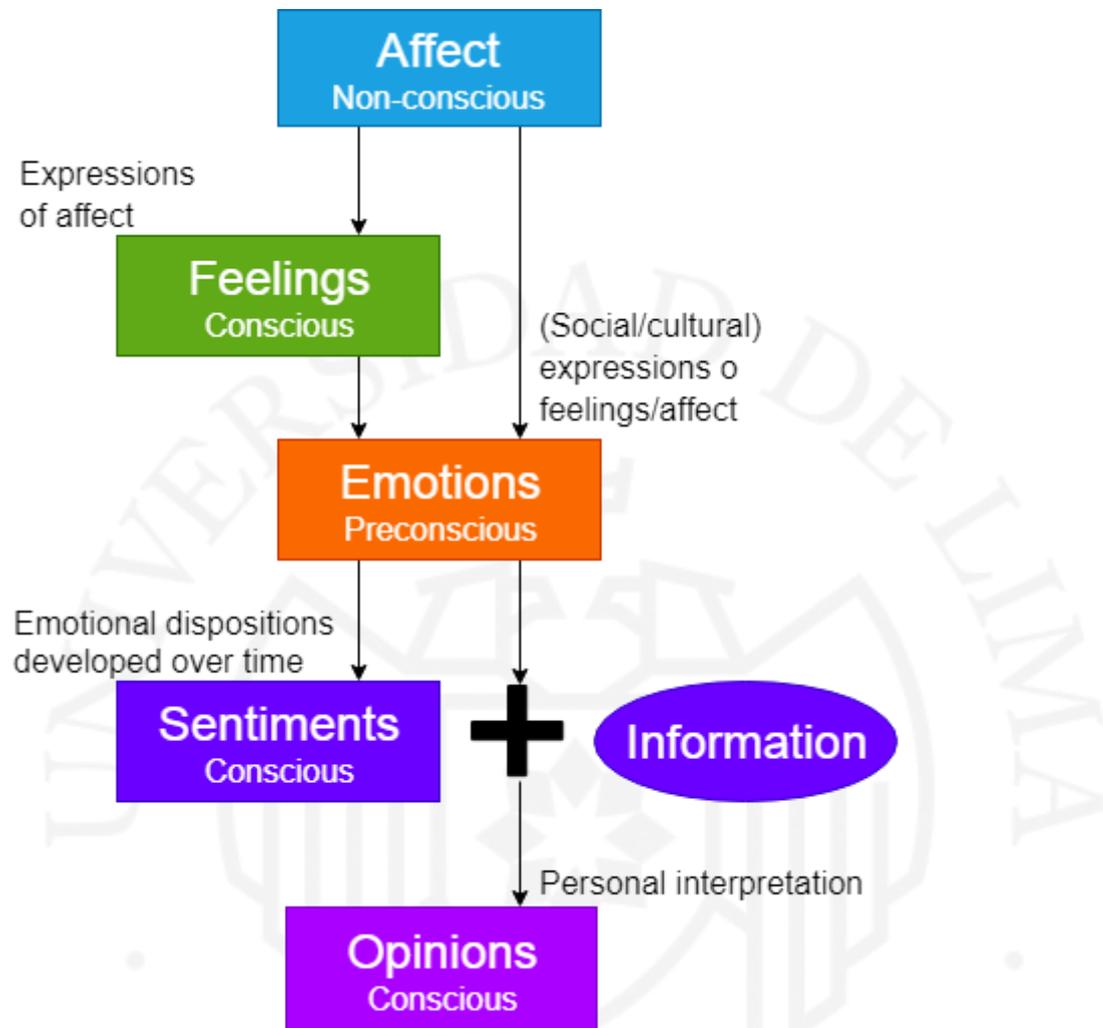
Mulcrone (2012) señala que NLP, al estar interesado en la detección automatizada de emociones a través del análisis de textos, emplea en la base de sus aplicaciones metodologías creadas por campos de estudio que ayudan a lograr discernir la subjetividad de la objetividad, uno de ellos es el Análisis de Sentimiento (Sentiment Analysis o SA).

De acuerdo a Cambria (2016) el SA contribuye en el crecimiento de la IA y aquellos campos de investigación que derivan de ella. Sostiene también que es un componente de gran valor para otros sistemas, ya que “la comunicación y aprendizaje entre humanos dependen de las emociones” (Cambria, 2016).

Por otro lado, en investigaciones de NLP, respecto al análisis textual, se suele tratar como similares los siguientes elementos subjetivos: afecto, sensación, emoción, sentimiento y opinión (Munezero, Suero Montero, Sutinen, y Pajunen, 2014), no obstante, difieren entre ellos (Figura 2.1). Afecto es el conjunto de expresiones observables de una emoción que ha sido experimentado y es sencillo de conceptualizarlo en el texto; sensación, es el conjunto de fenómenos conscientes que han sido etiquetados y pueden ser detectados a partir del texto; emoción, es el fenómeno psicológico complejo que es casi imposible de detectar en la totalidad del texto; sentimiento, es el fenómeno que soporta las disposiciones emocionales que se han desarrollado con el tiempo sobre objetos particulares; finalmente, opinión, es el conjunto de interpretaciones personales de información que pueden o no estar asociados a una emoción o sentimiento (Munezero et al., 2014).

Figura 2.1

Relación y diferencias entre elementos subjetivos en texto.



Nota: Modelo de relación, jerarquía y diferencias entre afecto, sensación, emoción, sentimiento y opinión.

Fuente: Munezero, M. et al. (2014)

Por ello, es importante tener noción de estos conceptos y sus diferencias con la finalidad de establecer el enfoque de una investigación o diseño de un sistema con relación a estos temas, y así aplicar los métodos que mejor se acoplen al cumplimiento de objetivos trazados por los investigadores. Por ejemplo, se aplica SA en reseñas y comentarios de un restaurante para analizar la satisfacción de sus comensales, pero también podría aplicarse Minería de Opiniones (Opinion Mining) o mecanismos que sinteticen las opiniones para luego determinar si son positivas, negativas o neutras.

Por otra parte, en los estudios de NLP relacionados al análisis textual, es importante definir de qué manera se procesará el texto que contiene estos elementos subjetivos, como por ejemplo, las emociones. Ante ello, Shirvahe, Garg y Mishra (2015) explican sobre métodos para reconocer en un texto la presencia de estos elementos; utilizan el método Keyword Spotting Technique (KST), que consiste en desagregar el texto a nivel de palabras, las cuales son convertidas en tokens, para luego compararlas con palabras clave predefinidas, relacionadas a una emoción, y evaluar si son la misma palabra o sinónimas. Seguido, se analiza la intensidad emocional que estas palabras identificadas presentan y se revisa el texto a nivel de oración para detectar el nivel de negación en ella. Finalmente, se determina cuál es la emoción dominante en todo el texto (Shirvahe et al., 2015).

Asimismo, frente a la necesidad de analizar en el texto la presencia de emociones, se llevaron a cabo estudios en los cuales los investigadores indicaron sus propuestas de solución y sus resultados. A continuación, se describen tres artículos sobre arquitecturas de sistemas para el reconocimiento y análisis de emociones y estados de ánimo en texto; la primera publicada en la Conferencia Internacional de Computación, Comunicación y Automatización (ICCCA) en 2015 (Shirvahe et al., 2015), y las dos siguientes en la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN) (Francisco y Gervás, 2006; García y Alías, 2008), respectivamente. En estas arquitecturas se sigue un modelo cuya entrada es un texto plano, y su salida es la atribución de una emoción hacia este texto (Francisco y Gervás, 2006; García y Alías, 2008; Shirvahe et al., 2015).

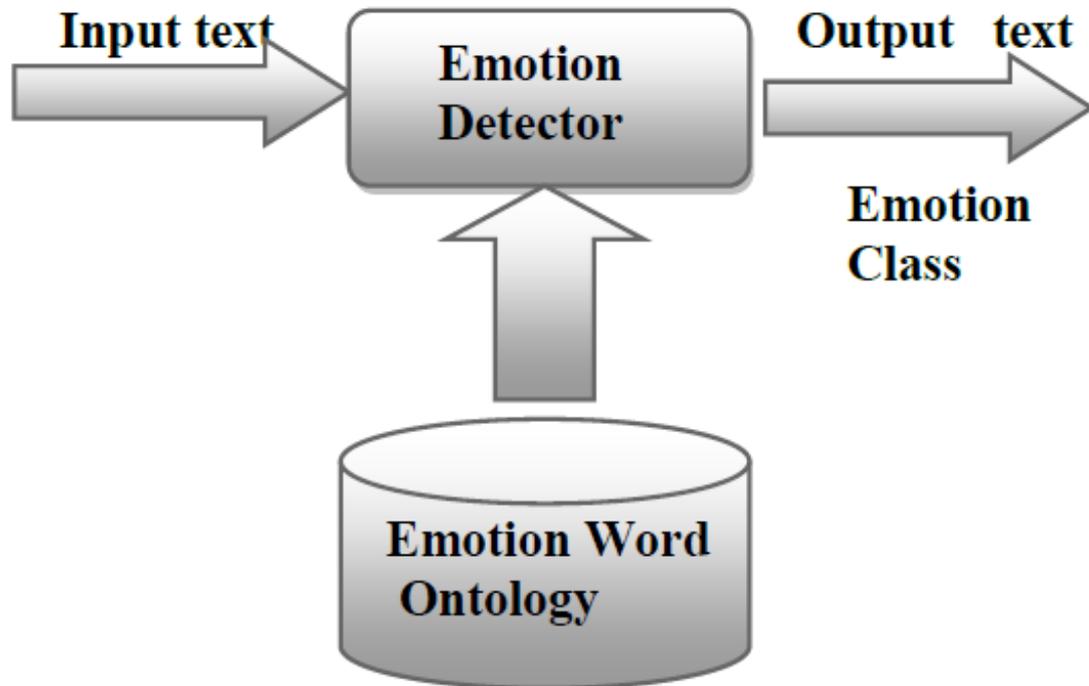
La primera arquitectura (Shirvahe et al., 2015) se modela de acuerdo a la manipulación de palabras clave y del concepto de Ontología; por ello, emplea KST para hallar emociones en texto (Figura 2.2). Como data de entrada o input, recibe la parte de un documento de texto o de un blog en inglés, el cual es procesado por este sistema compuesto de dos módulos: un Detector de Emociones, el cual funciona sobre la base de un algoritmo propuesto por los autores y una Ontología Emocional, la cual es una base de datos léxica que contiene aspectos específicos de cada palabra y las distribuye en dominios de acuerdo a su significado (Ávila, 2016). En este trabajo emplean ontologías para incrementar la eficiencia del proceso de categorización de textos, ya que “le permite a un programador especificar, de forma abierta y significativa, los

conceptos y las relaciones que caracterizan colectivamente a un dominio” (Shirvahe et al., 2015). Por consiguiente, para este escenario, su universo equivale a todas las palabras y cada dominio corresponde a un tipo de emoción, cuyos componentes son palabras relacionadas a ella (como sinónimos o los adjetivos que hacen una referencia a esa emoción). Al no haber, de acuerdo a Shirvahe et al. (2015), un modelo preestablecido de jerarquía de emociones ni tampoco un patrón estándar de relación entre palabras y emociones, es posible crear uno o varios modelos a su manera. Es así que, en esta investigación, la clasificación que desarrollaron fue en función al dominio de la psicología cognitiva. Por eso, establecieron que el objeto base para todos los elementos de la ontología es la palabra ‘eso’ o ‘algo’; a partir de ella, las palabras que representan una emoción, y también las clases de emociones según lo delimitado en la investigación, son entidades hijas, pues dependen de la clase padre para poder definirse, por ejemplo, ‘alegría’ es ‘algo’ y ‘enojo’ es ‘eso’ (Shirvahe et al., 2015).

El módulo Detector de Emoción contiene la mayor parte de la lógica del sistema porque en él se encuentra un algoritmo, desarrollado por los mismos investigadores (Shirvahe et al., 2015), cuya función es consultar las palabras encontradas en la ontología y ver si existe alguna relación entre las palabras claves del texto con la base de datos lingüístico emocional; en caso exista lo cataloga a dicha emoción y vuelve a analizar el texto, pero esta vez a nivel semántico. Adicionalmente, este algoritmo trabaja junto con la librería o API Apache Jena (The Apache Software Foundation, 2010) para interrelacionar la ontología y el aplicativo desarrollado en Java (Oracle Corporation, 1996).

Figura 2.2

Arquitectura N° 1 para sistema de detección de emociones en texto.



Nota: Arquitectura de sistema de detección de emociones en un texto.

Fuente: Shivhare, S. et al. (2015)

Para llevar a cabo su labor, este algoritmo debe cumplir con las siguientes características:

- Herencia. Esta medida también es un atributo importante en la solución porque en la realidad existen múltiples matices de emociones y combinaciones entre ellos, y esto también se plasma en el discurso de una persona. Por ello, para poder atribuirle una etiqueta emocional, Shivhare, Garg y Mishra (2015) establecieron niveles para organizar su ontología. Son seis sus emociones primarias: amor, tristeza, alegría, sorpresa, miedo y enojo. Al establecerse esta jerarquía de palabras y emoción por niveles en la ontología, se facilita la cuantificación de las veces que fue empleada una palabra correspondiente a una emoción, la búsqueda de estas con su valor numérico en la ontología, en conjunto con Apache Jena (The Apache Software Foundation, 2010), y se facilita el cálculo final. Por lo tanto, el puntaje de la palabra emotiva clave hallada en el texto perteneciente al nivel secundario o terciario de la Ontología Emocional, afectará al que se encuentra en el primer nivel (Shirvahe et al., 2015).

- Profundidad en la Ontología. Este atributo propio de la base de datos ayuda a señalar a qué nivel corresponde el nodo (palabra clave emotiva), ya que da una idea de qué tan concreta es dicha palabra con respecto a su estructura ontológica correspondiente. A mayor especificidad, menor profundidad (Shirvahe et al., 2015).
- Frecuencia en el texto. Las veces que aparece dicha palabra emotiva en el documento influye en el puntaje correspondiente a su dominio en la ontología, ya que le da mayor importancia a esa clase de emoción (Shirvahe et al., 2015).

En el algoritmo de Shirvahe et al. (2015), los elementos denominados Nodes[Ontology] denotan clases de emoción (palabras emotivas) en la ontología; Parent [j], clases padres en la ontología; Child [j], clases hijas en la ontología; Freq [m] indica frecuencia de la clase m^{th} en el documento de texto; Depth denota la profundidad de una clase particular en la ontología de la emoción a partir de la clase raíz; y Score[parent], la puntuación de la clase principal en la ontología. Tener en cuenta que esta calificación resultante es directamente proporcional a la frecuencia con la que se ha utilizado el término, e inversamente proporcional a su profundidad en la ontología (Algoritmo 1). Esta fórmula calcula la puntuación para cada clase de emoción de nivel primario disponible en la "Ontología Emocional" sumando los puntajes de las clases de emoción de sus niveles secundario y terciario respectivos. Después de calcular los puntajes para cada de emoción de nivel primario, los compara hasta hallar la mayor puntuación; finalmente, aquella clase de emoción detectada con el mayor puntaje se declara la emoción resultante del documento procesado (Shirvahe et al., 2015).

Algoritmo 2.1

Algoritmo para determinar el valor de una emoción en un texto

```
for j ← 1 to No. of Nodes[Ontology]
  do parent[j] ← parent of node j
  child[j] ← child of node j
  for m ← 1 to No. of Nodes[Ontology]
    do freq[m] ← frequency of occurrence of mth
    depth[m] ← depth of mth node in ontology
```

Calculate (x):

```
for m ← 1 to No. of Nodes[Ontology]
  score(x) ← freq[root]/depth[root]
  for m ← 1 to No. of parent nodes[Ontology]
    score(parent) = score(parent) + score(child)
  return score(parent)
for m ← 1 to No. of parent nodes[Ontology]
  emotion class ← High score[parent]
  return emotion class
```

Nota: Algoritmo desarrollado para determinar el valor de una emoción en un texto.
Fuente: Shivhare, S. et al. (2015)
Elaboración propia.

La segunda arquitectura utiliza los diccionarios afectivos ANEW (Bradley y Lang, 1999) y ANSW (Redondo, Fraga, Padrón, y Comesaña, 2007), y emplea KST - por su practicidad de implementación- para captar las palabras que indican el valor emocional del texto; en consecuencia, el costo computacional que implica es mínimo (García y Alías, 2008). Está estructurado en base al modelo psicológico emocional ‘Dimensiones Emocionales’ (Schlosberg, como se citó en García y Alías, 2008) debido, principalmente, a su compatibilidad con los diccionarios afectivos.

García y Alías (2008) modelaron esta arquitectura en base a las propuestas de Francisco, Hervás, y Gervás (2005), así como las de Díaz-Agudo, Gervás, Hervás, y Peinado (2004). Además, la construyeron para procesar textos de cualquier idioma; para demostrarlo, realizaron pruebas con textos en inglés y en español. Lamentablemente,

tuvieron percances durante el desarrollo para el análisis en español, por lo cual su sistema procesa solo textos en inglés (García y Alías, 2008).

Esta arquitectura de García y Alías (2008) ha sido implementada en Java (Oracle Corporation, 1996), y está compuesta de 8 módulos, los cuales se encargan de realizar un análisis léxico, luego sintáctico-semántico, y finalmente contextual. Asimismo, cuenta con seis emociones primarias que determinan la emoción resultante del texto: tristeza, alegría, sorpresa, miedo, enojo y ‘estado neutro’; esta última categoría indica si el documento o blog no presenta ninguna emoción (Figura 2.3).

Los dos primeros módulos, ‘Analizador Lexicográfico’ y ‘POS Tagger’ se encargan de dividir el texto en palabras, con el fin de obtener su categoría gramatical – si es un verbo, un sustantivo, un artículo, entre otros– para poder “filtrar aquellas sin significado afectivo potencial (stop words), como por ejemplo, los artículos” (García y Alías, 2008). Las palabras cuya categoría gramatical es sustantivo, adjetivo o verbo pasan al módulo ‘Desambiguador Semántico’.

El tercer módulo es un desambiguador semántico cuyo objetivo es reducir la ambigüedad del significado de los sustantivos del texto, según la frase a la cual pertenece (García y Alías, 2008). Una vez culminada la desambiguación, los tokens determinados y sus correspondientes significados pasan al módulo siguiente.

El cuarto módulo, ‘Búsqueda de Synsets’, cumple la función de hallar los sinónimos, antónimos e hiperónimos de los verbos, sustantivos y adjetivos del texto, según su significado previamente hallado (García y Alías, 2008). De esta manera, si no se encuentra una de estas palabras en los diccionarios afectivos, se puede buscar utilizando su sinónimo, antónimo o hiperónimo. Así, la capacidad de búsqueda del sistema se amplifica.

El quinto módulo, ‘Flexionador’, se encarga de lematizar, es decir, apartar la flexión –los prefijos y los sufijos- de las palabras potencialmente afectivas, así como las de sus sinónimos, antónimos e hiperónimos, para quedar solamente con los lexemas o raíces y luego buscarlas en el lexicón. Con ello, se logra mayor abarque de búsqueda del idioma; es así que este módulo “resulta fundamental para los idiomas fuertemente

flexionadores como el español” (García y Alías, 2008). Por ello, para los textos en español, utilizan la librería Snowball (Porter, MacFarlane, y Boulton, 2002).

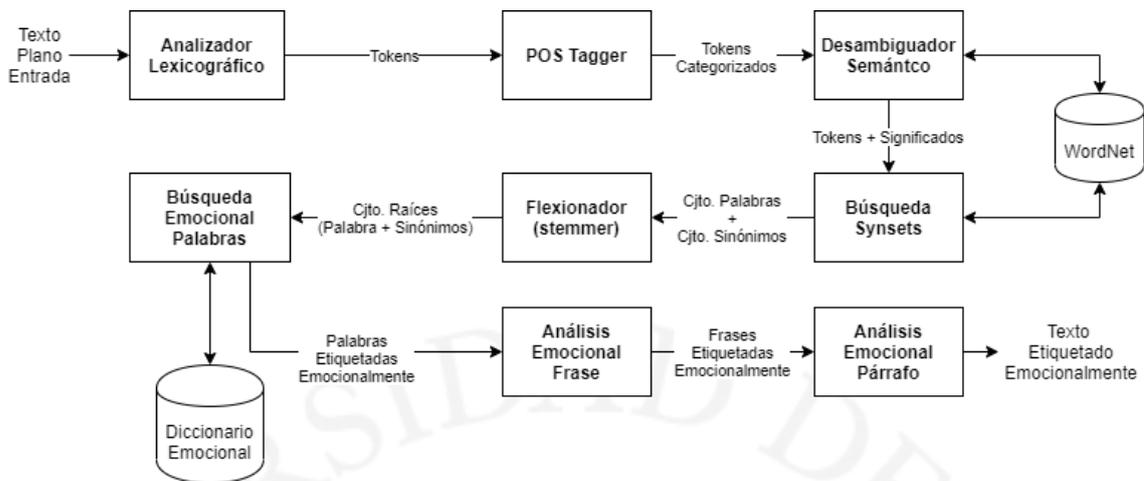
Al sexto módulo, ‘Búsqueda Emocional de palabras’, llegan las palabras potencialmente afectivas ya lematizadas –así como sus sinónimos, antónimos e hiperónimos– para proceder con la búsqueda de estas en el diccionario afectivo según su lexema y su categoría gramatical. Previamente, en este diccionario afectivo, las palabras que lo conforman, y su correspondiente categoría gramatical, han sido calificados de 0 a 10 en tres dimensiones: control, activación y valencia (García y Alías, 2008). En este módulo, se procede con la categorización emocional de las palabras filtradas del texto: si es positivo el resultado de la búsqueda usando la palabra o su sinónimo, los valores correspondientes a valencia, activación y control del diccionario se le atribuyen a la palabra del texto; si es positiva la búsqueda usando el antónimo, se atribuyen los valores complementarios; si es negativo en ambos casos de búsqueda, el valor asignado a esa palabra es -1 (García y Alías, 2008).

Finalmente asignada las magnitudes respectivas a las palabras detectadas, los dos últimos módulos, ‘Análisis Emocional Frase’ y ‘Análisis Emocional Párrafo’, calculan el promedio de los valores de las palabras emocionales que la componen, el cual determina la emoción resultante o etiqueta emocional de la oración; se reitera esta operación entre frases para categorizar los párrafos del texto y, por último, se repite este procedimiento entre párrafos para determinar la emoción resultante del texto (García y Alías, 2008).

En síntesis, el sistema divide el texto en palabras para encontrar aquellas relacionadas a una emoción; luego, las procesa y contabiliza sus valores numéricos hallados en el diccionario afectivo; después, calcula el promedio de cada clase de emoción a nivel de frase, de párrafo y de texto; y finalmente, aquella categoría que tiene mayor promedio es la emoción resultante del texto (Figura 2.3).

Figura 2.3

Arquitectura N° 2 para análisis y detección de emociones a partir de un artículo



Nota: Arquitectura de detección y análisis de emociones a partir de un artículo.

Fuente: García, D. y Alías, F. (2008)

Elaboración propia.

La tercera arquitectura, elaborada por Francisco y Gervás (2006) para cuentos en inglés, se caracteriza por su capacidad de análisis ante la presencia de negación en las oraciones. Para lograrlo, fusionaron dos técnicas de categorización de emociones en texto: Afinidad léxica y Keyword Spotting Technique (KST); el motivo para hacerlo fue “tratar de superar las desventajas de ambas mediante su combinación” (Francisco y Gervás, 2006). En Tabla 2.1 se puede observar estas desventajas.

Tabla 2.1

Desventajas de las técnicas KST y Afinidad Léxica para categorizar textos en una emoción

KST	Afinidad Léxica
Inconvenientes ante la presencia de la negación en la frase.	No distingue el contexto en el cual se encuentra la palabra.
Dependencia del método en aspectos superficiales (p. ej. establecer contenido emocional según los adjetivos usados).	Dificulta el desarrollo de un modelo reutilizable e independiente, ya que suele entrenarse con corpus.

Fuente: Francisco, V. y Gervás, P. (2006)

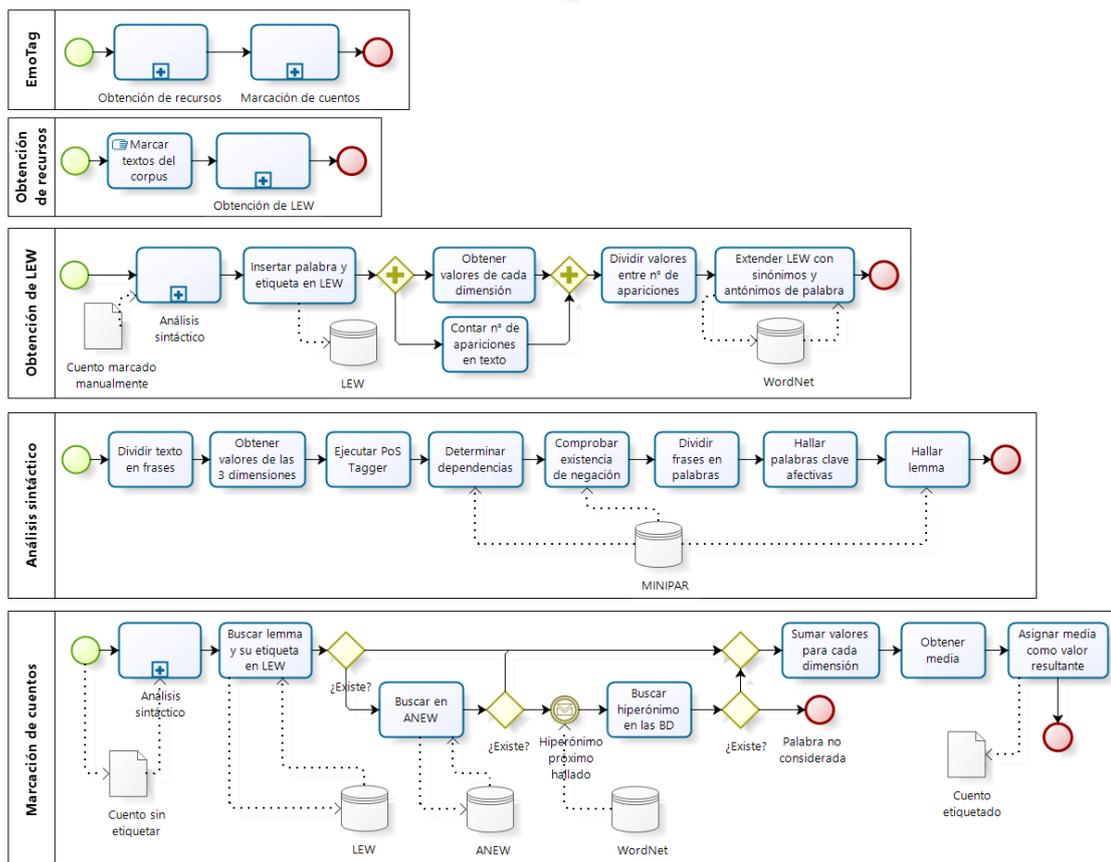
Elaboración propia.

La arquitectura está compuesta de dos grandes procesos (Figura 2.4): el primero consiste en obtener los principales materiales para la categorización emocional, que son

“un corpus de cuentos con frases anotadas emocionalmente y una lista de palabras (LEW [List of Emotional Words]) con su correspondiente relación con cada una de las dimensiones emocionales” (Francisco y Gervás, 2006); el segundo proceso consiste en ejecutar este marcador emocional en cualquier cuento, utilizando los valores de las tres dimensiones emocionales.

Figura 2.4

Arquitectura N° 3 para detectar emociones en un texto



Nota: Diagrama de procesos de la tercera arquitectura, la cual consiste en primero construir los recursos necesarios para el sistema –el corpus de cuentos con etiquetas emocionales y la lista de palabras relacionadas a emociones–, para después analizar sintácticamente el texto, y por último etiquetar el cuento con su emoción resultante.

Fuente: Francisco, V. y Gervás, P. (2006)

Elaboración propia.

Además de la combinación de técnicas, implementaron un módulo de análisis de dependencia para hallar la localización de aquellas palabras que generaban negación o eran afectadas por la negación (Francisco y Gervás, 2006). Este módulo genera el árbol de dependencias de cada palabra, proporcionando de ellas toda la información

pertinente para armar la jerarquía sintáctica del texto, determinar la herencia entre palabras, y organizar información adicional, como el análisis léxico, por ejemplo.

En síntesis, estas tres arquitecturas (Francisco y Gervás, 2006; García y Alías, 2008; Shirvahe et al., 2015) comparten las siguientes características:

- Emplean la técnica KST para detectar emociones en el texto y clasificarlo a una emoción.
- Sus entradas son fragmentos textuales, pertenecientes a un blog o un documento. No trabajan con párrafos redactados en tiempo real por una persona.
- Para el análisis del texto, lo dividen en frases primero para procesar cada palabra que lo compone; en otras palabras, consideran la oración como la unidad atómica del escrito. Después, se examina a nivel de párrafo, y por último, a nivel de todo el documento.
- Han sido desarrolladas y probadas para textos en inglés.
- Emplean una base de datos léxica emotiva, en la cual previamente se han clasificado las palabras en emociones, y también se establecen cuáles son sus valores numéricos.

De las características mencionadas, es relevante destacar la importancia de la base de datos léxica emotiva porque la actualización de los múltiples diccionarios y bases de datos léxicas va a una velocidad menor que los cambios lingüísticos dados en el día a día, así como también solo abarca el uso estándar de la lengua, de tal forma que excluye expresiones que forman parte del lenguaje familiar y coloquial, así como las manifestaciones lingüísticas de la mayoría de los grupos sociales, de las cuales algunas han pasado al dominio común, como las jergas y los argots (Santamaría, 2007). Por ejemplo, la base de datos léxica generada por expertos en análisis de opiniones de la empresa Autoritas Consulting, la cual está conformada por 1300 palabras positivas y 1270 palabras negativas en idioma español (Rangel, 2016).

CAPÍTULO III: MARCO TEÓRICO

3.1 Estado de ánimo

De acuerdo a Thayer (1998), es un estado psicológico que se ve influenciado y va a la mano de un conglomerado de sensaciones autopercibidas (cuyos orígenes son de naturaleza biopsicológica y prácticamente manejables por el propio sujeto perceptor) que determinan nuestro bienestar psicológico, la calidad de las relaciones interpersonales y qué tan eficaz puede uno comportarse.

André (2010) brinda una visión general del estado de ánimo:

Podríamos definir los estados de ánimo diciendo que son contenidos mentales, conscientes o inconscientes, que mezclan estados corporales, emociones sutiles y pensamientos automáticos, y que influyen en la mayoría de nuestras actitudes. En general les prestamos escasa atención, y tampoco les consagramos esfuerzos para intentar comprenderlos, integrarlos en nuestra reflexión, o pedirles que se pongan a nuestro servicio. Por fortuna, lo hacen por sí mismos, solitos: su papel y su influencia sobre lo que somos y hacemos somos inmensos. (p. 20)

La modificación de este se debe a factores biológicos, como el sueño, la salud, el estado físico, los ciclos diarios (día y noche, por ejemplo), el ejercicio y la nutrición; pero también afectan variables como la cognición y el estrés, así como los sucesos que presentan un impacto en la vida cotidiana. En resumen, “el estado de ánimo es algo así como un termómetro clínico que refleja todos los acontecimientos externos e internos que nos afectan” (Thayer, 1998).

Según las afirmaciones de Paul Ekman al Dalai Lama y David Goleman (Goleman, 2003), existen dos formas para producir un estado de ánimo: a partir de cambios internos cuya relación con los sucesos en el entorno es casi nula, y a raíz de experiencias emocionales muy intensas. Así, se confirma que los estados de ánimo pueden ser producto de factores internos o biológicos, o externos cuya relación con el ambiente es estrecha.

Reeve (2010) se fundamenta en teorías evolutivas e investigaciones neurocientíficas para afirmar que existen dos tipos de estados de ánimo: positivos y negativos. Los estados de ánimo positivos provocan actitudes de aproximación hacia su entorno, como interés, curiosidad y entusiasmo. Asimismo, esclarecen la mente de la persona en el planteamiento de objetivos; permiten una visión y procesamiento de información más rápida, global e intuitiva; así como también mejoran el autocontrol, la memorización de información trascendental y la capacidad de persuasión (André, 2010). Adicionalmente, promueven el deseo por brindar ayuda, básicamente porque le genera a la persona la sensación de seguridad y confianza de que su entorno no es peligroso, y porque logra que una persona se sienta conforme consigo misma luego de dar apoyo (Isen, como se citó en Jhangiani y Tarry, 2014).

Por otro lado, los estados de ánimo negativos conducen a tratar la información que capta de su medio ambiente de manera pausada, atenta y detallada, dando así la impresión de que el tiempo pasa más lento; también provocan actitudes de evitación hacia su entorno, como el desdén, la irritación y el retraimiento. Además, dirigen la atención hacia aquello que podría parecerles peligroso; generan debilidad ante los cambios de estados de ánimo; abocan a la toma de decisiones que superen las capacidades de la persona; incitan a realizar acciones y consumir productos y servicios que prometen bienestar inmediato –como despilfarrar dinero o comer y beber en exceso– para luego, por la culpa y la inconformidad consigo mismo, decidir por autocastigarse en lugar de autorrepararse, derivando así en un círculo vicioso (André, 2010). Sin embargo, con el fin de aliviar el malestar generado por los estados de ánimo negativos, se incrementa el deseo de crear buenas relaciones con otras personas (Jhangiani y Tarry, 2014), dando a producirse el efecto Macbeth: “las personas tienden a querer limpiarse cuando perciben que han violado sus normas éticas” (Zhong y Liljenquist, como se citó en Jhangiani y Tarry, 2014).

No obstante, las actitudes provocadas según el tipo de estado de ánimo hacia su entorno también pueden ser producto de la interacción con personas o de reacciones frente a diversas situaciones. Como señala André (2010):

Si el acercamiento es gratificante o tiene el éxito como resultado, entonces se siente alegría, placer, buen humor, que facilitarían otros acercamientos. Pero si

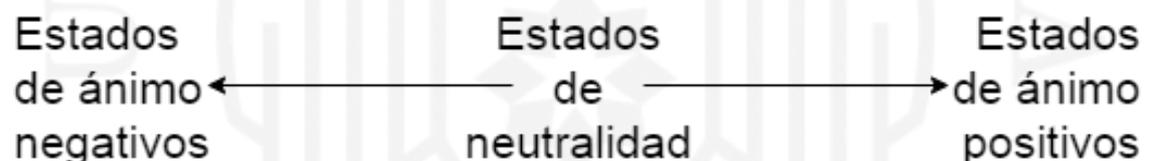
resulta que su resultado es un fracaso o disgusto, uno se llena de tristeza, decepción, resentimiento y aversión, que comprometerán o frenarán otros acercamientos. (p. 53)

Es importante resaltar lo propuesto por André (2010), quien señala que es errónea la relación inversamente proporcional entre estos dos tipos de estados de ánimo (Figura 3.1), pues el incremento de los estados de ánimo negativos en una persona no contrarresta sentir los estados de ánimo positivos, y viceversa. Incluso, esta visión plantea que, cuando ambos polos tienen el mismo nivel, se produce una neutralidad entre estados de ánimo. “Se ha comprobado que ciertos estados de ánimo, como la nostalgia o la tristeza, pueden ser mixtos” (André, 2010).

Figura 3.1

Visión errónea de la relación entre estados de ánimo positivos y negativos

Modelo unidimensional de los estados de ánimo



Nota: Modelo que refleja la visión que el comportamiento de los estados de ánimo se da de forma unidimensional, el cual es erróneo.

Fuente: André, C. (2010)

Elaboración propia.

En realidad, el comportamiento de los estados de ánimo se plasma de forma bidimensional porque reprimir estados de ánimo negativos no necesariamente equivale a que solo estén presentes estados de ánimo positivos y viceversa; de hecho, uno puede encontrarse en un estado en el cual se mezclan ambos tipos (André, 2010).

Un ejemplo que plantea este autor ilustra cómo la relación entre estados de ánimo no es lineal ni tampoco se produce esa neutralidad entre estados de ánimo, como aparece en el modelo anterior (Figura 3.1):

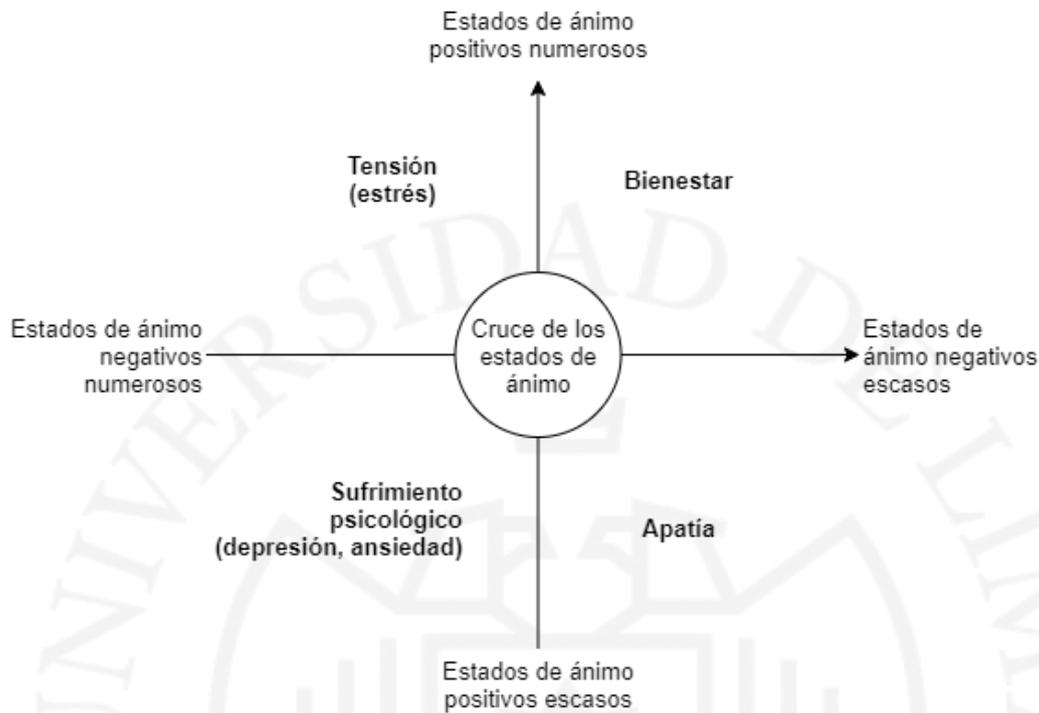
Cuando todo marcha bien materialmente a mi alrededor, no me siento estresado ni feliz, sino tranquilo (dicho lo cual, según los entornos en los que evolucionamos, ¡la tranquilidad puede ser un estado de ánimo que ya tiende hacia lo positivo!). Reprimir los estados de ánimo negativos cuando se nos dice: «No llores», no conduce a estados de ánimo positivos, ni hace que brote en nosotros una sonrisa. Igualmente, cuando decimos: «No está nada mal», o: «No tengo de qué quejarme», nuestro discurso significa que no nos asaltan estados de ánimo negativos, pero que tampoco estamos sumergidos en estados de ánimo positivos. Se trata de tranquilidad, no de felicidad. (p. 58)

La Figura 3.2 muestra que los estados de ánimo dependen de la polaridad y de la intensidad, en comparación con el modelo anterior que solo depende de la polaridad. El punto de origen es el cruce de los estados de ánimo y partiendo de ella, de acuerdo a la cantidad presente de cada tipo de estado de ánimo, existen cuatro estados: tensión, sufrimiento, apatía y bienestar.

Figura 3.2

Modelo bidimensional de los estados de ánimo.

Modelo bidimensional de los estados de ánimo



Nota: Modelo bidimensional del comportamiento de los estados de ánimo en una persona. A mayor presencia de estados de ánimo positivo y negativo, se produce Tensión, vinculada a estrés, porque esta mezcla genera efervescencia y cansancio. Ante una alta presencia de estados de ánimo negativos y baja presencia de estados de ánimo positivos, se da Sufrimiento psicológico, denominado también ‘depresión’ o ‘ansiedad’. Cuando es escasa la presencia de estados de ánimo negativo y positivo, se genera Apatía. Por último, al haber un alto nivel de estados de ánimo positivos y bajo nivel de estados de ánimo negativos, se produce Bienestar.

Fuente: André, C. (2010)

Elaboración propia.

Cabe mencionar que la transición del estado de ánimo de un tipo a otro varía en qué tan fuerte se da este cambio (Silver, 2018). De hecho, es relativamente normal, natural y habitual que una persona experimente cambios de estado de ánimo en situaciones que lo exaltan positiva o negativamente (Kilô, 2017; Rovira, 2018). Por ejemplo, pensar haber desaprobado un examen cuando finalmente se ha logrado aprobarlo con la máxima puntuación; o haber culminado un reporte importantísimo en el equipo cuando de pronto se va la luz y no se ha guardado el archivo.

No obstante, aunque estos cambios de estado de ánimo –“mood swing” en inglés– son parte del día a día de una persona, le es trabajoso sobrellevarlos porque se

encuentra en un estado de ánimo de angustia, el cual es producto de “la impaciencia, la ansiedad, el estrés y la fea costumbre de preocuparse por todo” (Kilô, 2017). Estas fluctuaciones son peligrosas si son intensas u ocurren frecuentemente, pues llegan a interferir con la manera de cómo vive la persona a un nivel que lo perjudique (Schimelpfening, 2018). Incluso, de acuerdo a Rincón (2013), el enojo y la tristeza afectan negativamente el estado de ánimo de una persona, pues desencadenan eventos que arriesgan su vida; estos también propician actitudes negativas como apatía, desmotivación y bajo rendimiento (Rovira, 2018). En otras palabras, como repercute en el pensar y actuar de una persona, la alteración del estado de ánimo en una alta intensidad y frecuencia produce consecuencias que dañan su salud y su relación con su entorno (Catalán, 2017). Por ello, es importante detectar aquello que afecte el estado de ánimo de una persona y determinar el grado en que lo hace (Silver, 2018).

3.2 Emoción

Las emociones “son fenómenos de corta duración, relacionados con sentimientos, estimulación, intención y expresión, que nos ayudan a adaptarnos a las oportunidades y retos que enfrentamos durante los sucesos significativos de la vida” (Reeve, 2010).

Kleinginna y Kleinginna (como se citó en Munezero et al., 2014) poseen otra definición sobre las emociones:

[Las emociones son] un conjunto complejo de interacciones entre los factores subjetivos y objetivos, mediada por los nervios y los sistemas hormonales, lo que puede a) dar lugar a experiencias afectivas, tales como sentimientos de excitación, el placer y el disgusto; b) generar procesos cognitivos como la percepción emocional relevante afectar, tasaciones, los procesos de etiquetado; c) activos generalizadas ajustes fisiológicos a las condiciones excitantes; y d) llevar a un comportamiento que es a menudo, pero no siempre expresiva, dirigido a un objetivo y adaptativa. (p. 102)

En cambio, Shouse (como se citó en Munezero et al., 2014):

[Las emociones] son las proyecciones o pantalla de un sentimiento y, a diferencia de los sentimientos, la visualización de las emociones puede ser verdadera o fingida. [...] nosotros transmitimos emoción hacia el mundo; a veces esa emisión es una expresión de nuestro estado interno y otras veces es artificial con el fin de cumplir con las expectativas sociales. (p. 103)

Scherer (como se citó en Munezero et al., 2014) presenta una definición similar a Reeve (2010), pero equilibra los factores cognitivos como biológicos:

Episodios de cambios coordinados en varios componentes (incluyendo al menos la activación neurofisiológica, la expresión del motor y la sensación subjetiva, pero posiblemente también tendencias a la acción y los procesos cognitivos) en respuesta a eventos externos o internos de gran importancia para el organismo. (p. 102)

Es importante tener en cuenta que la intensidad de una emoción se ve influenciada por el proceso de valoración. En general, este proceso consiste en que las personas juzgan si un suceso tiene relevancia personal para su bienestar físico, psicológico, metas, estado financiero o relaciones interpersonales; pero también las capacidades percibidas de afrontamiento influyen en esta evaluación, porque si uno puede encararlo, la importancia de esta situación es mínima. En conclusión, “las personas primero valoran su relación con el suceso vital (valoración primaria) y después valoran su potencial de afrontamiento dentro del mismo suceso (valoración secundaria)” (Reeve, 2010).

También funciona como indicador del estado de satisfacción de los motivos de una persona y su capacidad de adaptación: si logra conectar su motivación con sus estados emocionales, además de mostrar una adaptación exitosa, presentará emociones positivas (Reeve, 2010).

Para su estudio e investigación, se debe tomar en cuenta dos perspectivas: biológica y cognitiva. De acuerdo a la primera perspectiva, se dividen las emociones entre primarias (conformadas hasta por 10 emociones) y secundarias (equivalentes a las restantes) (Reeve, 2010). Para la segunda perspectiva, una emoción se evoca a partir de

procesos de valoración, y las teorías propuestas difieren entre sí fundamentalmente por las evaluaciones subjetivas de cada ser o a valoraciones de relevancia personal de la situación, o del objeto, etc. (Pérez y Redondo, 2006). No obstante, cada autor, de acuerdo a su perspectiva, tiene su idea de clasificación de las emociones; es decir, no existe una categorización definitiva porque ninguna plasma la totalidad del ámbito emocional humano (Corbin, 2017). Una muestra de ello es la propuesta de Corbin (2017), la cual abarca la mayoría de las teorías sobre la emoción y consta de ocho grupos: Primarias o básicas, Secundarias, Positivas o Saludables, Negativas o Tóxicas, Ambiguas o Neutras, Estáticas, Sociales e Instrumentales.

Figura 3.3

Clasificación de las emociones



Nota: Articulando las diversas teorías sobre las emociones, estas se clasifican en ocho grupos: primarias, secundarias, positivas, negativas, ambiguas, estáticas, sociales, e instrumentales.

Fuente: Corbin, J. (2017)

Elaboración propia.

3.2.1 Emoción y estado de ánimo

En el diálogo entre el Dalai Lama y Paul Ekman, este último indica que ambos conceptos difieren principalmente por la duración, porque “las emociones pueden aparecer y desaparecer en cuestión de minutos o de segundos, mientras que los estados de ánimo pueden llegar a durar varios días” (Goleman, 2003).

Goldsmith (citado en Reeve, 2010) indica que, de todos los criterios, tres son fundamentales para discernir emoción de estado de ánimo: antecedentes, especificidad de acción y curso temporal (Tabla 3.1).

Tabla 3.1

Diferencias entre emoción y estado de ánimo

Criterios	Emoción	Estado de Ánimo
Antecedentes	Deriva de situaciones vitales relevantes y de las estimaciones de su importancia para el bienestar del individuo.	Surge de procesos indefinidos y que a menudo se desconocen.
Especificidad de acción	Influye fundamentalmente en el comportamiento y dirigen cursos específicos de acción.	Repercuten en mayor medida en la cognición y dirigen aquello en lo que piensa la persona.
Curso temporal	Emana de eventos breves con duración de unos cuantos segundos o minutos.	Emana de sucesos mentales que duran horas o días.

Nota: La comparación está en función a los siguientes aspectos: Antecedentes, Especificidad de acción, y Curso temporal.

Fuente: Reeve J. (2010)

Elaboración propia.

La forma en cómo la emoción y el estado de ánimo interactúan es a través de la cotidianeidad: “la gente siente algo todo el tiempo. Lo que siente de manera típica es un estado de ánimo, un modo de sentirse que a menudo existe como efecto posterior a un episodio emocional experimentado anteriormente” (Reeve, 2010).

3.3 Natural Language Processing (NLP)

Es un campo de interés para la ciencia de la computación y lingüística debido a la interacción entre seres humanos y computadoras por medio del lenguaje; se trata de convertir la información de las bases de datos de los sistemas computacionales en un lenguaje entendible para la lectura de una persona (Kumar, 2011).

Para que un sistema NLP funcione correctamente, debe emular el comportamiento humano al nivel de entrada y salida, de lo contrario, no se lograría la comunicación deseable entre humano y máquina. La cuestión es si será posible imitarlo utilizando mecanismos computacionales, a pesar que la ciencia y psicología cognitiva indican qué sucede en la mente humana cuando procesa el lenguaje (Ball, 2006).

Aunque más allá de imitar conductas de personas, para hacer un trabajo en el área de NLP inclusive se debe tener noción sobre qué es lenguaje natural. ¿Por qué se añade este atributo: “natural”? De acuerdo a Polanco Fernández (2000), se le llama así porque no es artificial: no está construido en base a reglas de formación; o sea, bajo ningún control de alguna teoría o una gramática. Otra característica es la consistencia de sus oraciones: depende de la experiencia práctica humana; en otras palabras, se acomoda de acuerdo a cómo se manifiesta el lenguaje en ese instante. Para mantenerla, la combinación de palabras permitidas debe respetar la sintaxis y semántica establecidas en la población; es decir, las secuencias deben hacerse de forma correcta y con sentido, de tal manera que brinden orden al armar las oraciones que conforman el lenguaje (Polanco Fernández, 2000).

Es por ello que aún no se tiene definido un formalismo o mecanismo adecuado para representar el lenguaje natural; mucho menos, ya aterrizándolo en el ámbito científico, aquel que genere menor costo computacional: mientras más complejo sea el formalismo, menos eficiente será. No obstante, existen diversas propuestas brindadas por las escuelas lingüísticas generativas, las cuales también pueden complementarse entre sí. Entre ellas, están la escuela de Chomsky y sus cuatro clases fundamentales de lenguajes formales de acuerdo a la complejidad (lenguaje independiente de contexto); la Gramática Transformacional y la Gramática de Unificación (ambas, lenguaje dependiente de contexto) (Ramírez González, 2013).

Por lo tanto, de acuerdo a Carbonell (como se citó en Buenaga, 2001), el objetivo de esta tecnología es producir sistemas informáticos capaces de comprender o generar expresiones en lenguaje natural. En otras palabras, en computación se busca “poder especificar rigurosamente estos significados, por los métodos de interpretación de los sistemas formales. Estos métodos en cuestión, constituyen las semánticas del lenguaje formal” (Polanco Fernández, 2000).

El análisis al lenguaje en NLP, según Buenaga (2001), se divide en cinco niveles:

- Nivel fonético y fonológico. Consta de la concordancia entre sonidos y las palabras vocablos que los simbolizan. Se aplica principalmente en sistemas de reconocimiento y generación de voz (Buenaga, 2001).
- Nivel morfológico y léxico. Comprende el análisis de cada palabra como una unidad y cómo se pueden articular usando componentes más simples: los morfemas (Buenaga, 2001).
- Nivel sintáctico. Consiste en el análisis de cómo se conjugan palabras para armar oraciones; se examina el rol que puede atribuirse a cada palabra para estructurar sintagmas y oraciones (Buenaga, 2001).
- Nivel semántico. Consta de la investigación de los significados de cada palabra y cómo ensamblarlas para generar el significado de cada oración fuera de un contexto. Los sentidos que posee una palabra (sus significados de acuerdo a los diversos contextos) son las unidades base en este nivel, los cuales están organizados en un grupo de clases e objetos, que categorizan las palabras, denominado ontología (Buenaga, 2001).
- Nivel pragmático y de procesamiento del discurso. Consiste en la inspección de cómo aplicar las oraciones en función a un contexto y cómo influye en la interpretación de estas (Buenaga, 2001).

De acuerdo con Fernández (2011), se han desarrollado recursos para representar el lenguaje natural y así procesarlo a través de la informática. Entre estos recursos están las siguientes: bases de datos, léxicas, corpus, tesauros y las ontologías.

3.4 Sentiment Analysis (SA)

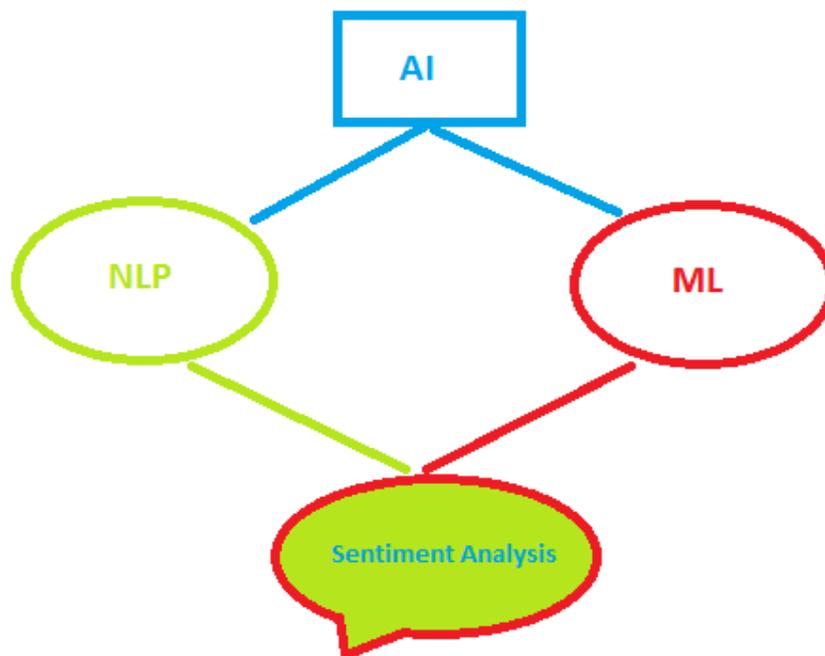
Sentiment Analysis –‘Análisis de Sentimiento’ en español- es “un campo de estudio que analiza las emociones, sentimientos, apreciaciones, opiniones, evaluaciones, y actitudes que las personas expresan hacia o sobre entidades como productos, servicios, organizaciones, individuos, eventos, problemas, temas, y los atributos de estos” (Liu, 2012).

Según Karimkhan (2017), para efectuar SA existen esencialmente tres enfoques. El primero se basa en léxico, es decir, solo se aplican los conceptos de NLP para determinar la polaridad del texto y, para ello, sugieren usar un diccionario afectivo. El segundo está basado en el aprendizaje automático: emplea NLP y Machine Learning (ML) para armar un modelo de clasificación y entrenarlo con data previamente seleccionada según su polaridad. El tercero es una mezcla de los dos enfoques anteriores y consiste en desarrollar un modelo de clasificación que trabaje con un diccionario afectivo y un banco de datos preseleccionados.

La definición de estos tres enfoques demuestra el vínculo de SA con NLP y con ML (Figura 3.4) (Karimkhan, 2017).

Figura 3.4

Relación de Análisis de Sentimiento con NLP y con Machine Learning



Nota: La matriz de los conceptos de Sentiment Analysis (SA), NLP y Machine Learning (ML) es la Inteligencia Artificial (IA). Los conceptos de NLP y ML influyen en la existencia de tres enfoques de la aplicación de Sentiment Analysis.

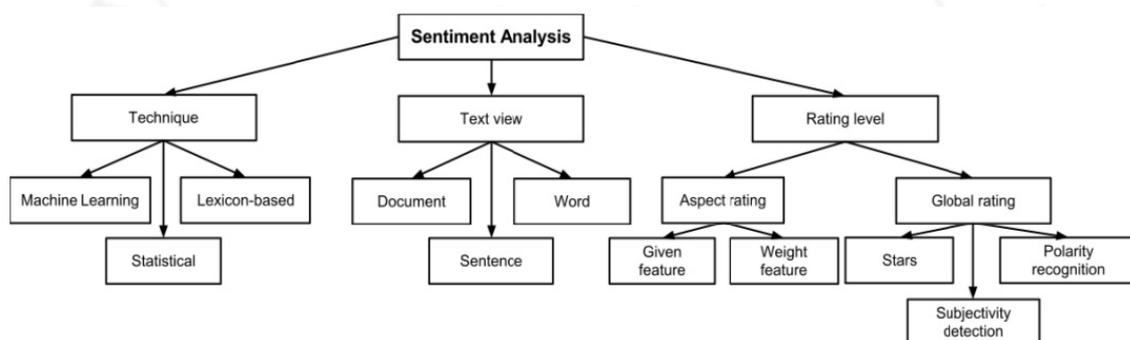
Fuente: Karimkhan, P. (2017)

Por otra parte, Collomb, Costea, Joyeux, Hasan y Brunie (2014) indican que los trabajos realizados en SA se puede clasificar principalmente según los siguientes criterios: la técnica de SA empleada, el nivel de detalle del análisis textual, el tipo de estructura del texto y el tipo de calificación.

Según la técnica de SA utilizada existen cuatro métodos para determinar el sentimiento resultante (Figura 3.5). El primero, basado solo en ML, consiste en aplicar algoritmos de aprendizaje y entrenarlos con data ya tratada. El segundo, basado solo en NLP, consiste en calcular la polaridad del sentimiento a través del significado de cada palabra u oración del texto analizado. El tercero, basado en modelos estadísticos, representa el texto analizado como una mezcla de puntuaciones y características. El cuarto, basado en reglas para categorizar sentimientos, consiste en hallar palabras que correspondan a un sentimiento y luego clasificarlas según su polaridad (Collomb et al., 2014).

Figura 3.5

Clasificación de Sentiment Analysis según el tipo de técnica empleada



Nota: Las clasificaciones más destacadas de Sentiment Analysis son según la técnica, cómo analiza el texto, y de qué manera se calificará el texto. La técnica está basada en el uso de un lexicon, de ML, o de estadísticas. En cuanto al texto, se puede procesar a nivel de documento, de oración, o de palabra. Fuente: Collomb, A. et al. (2014)

Según la estructura del texto, se determina la polaridad de sentimiento de acuerdo a uno de los tres niveles de análisis de sentimiento: palabra, oración o documento (Collomb et al., 2014).

3.5 Base de datos léxica

Es una colección de información lingüística. Se organiza de acuerdo a un modelo específico que permite el almacenamiento, recuperación y modificación de la data. En caso de seguir una estructura relacional o jerárquica, podrá guardar grandes volúmenes de datos atómicos. En la Lingüística, son empleadas como fuentes de información debido a su reutilización por otros recursos, como un lexicon computacional, por ejemplo (Fernández, 2011).

3.5.1 Tesoro

De acuerdo a Fernández (2011), la definición de tesoro es la siguiente:

Es un modelo que permite almacenar y recuperar información referente a un vocabulario controlado y dinámico de términos con relaciones semánticas entre ellos y que se aplican a un campo temático del conocimiento. Estas relaciones suelen ser de tres tipos fundamentales: equivalencia, asociativas y jerárquicas. (p. 14)

3.5.2 Corpus

Para Fernández (2011), el corpus está conformado por textos generados en eventos reales, los cuales son agregados si cumplen un conjunto de normas lingüísticas explícitas que garantizan que esta base de datos pueda emplearse como ejemplar característico de una lengua. Es decir, es un tipo de base de datos léxica que refleja el habla natural.

Asimismo, Oliver, Moré y Climent (2008) indican que el corpus “es un conjunto de datos lingüísticos que reflejan el uso de una lengua. [...] se puede obtener información sobre fenómenos que solo se pueden describir a partir de los datos y no desde un enfoque teórico”. Existen dos tipos: el corpus oral, que sirve para representar el habla y trabajar en el ámbito de la fonética; y el corpus escrito, en el cual se encuentran los corpus de referencia, los monitores y los paralelos (Oliver, Moré, y Climent, 2008).

Un corpus se caracteriza por su formato electrónico, pues sus textos que lo componen deben estar en dicho formato para realizar tareas automatizadas como localización de palabras o categorías gramaticales, recuperación de información registrada en él junto con su contenido anterior o posterior (conocido como ‘concordancia’), cálculo de la cantidad de veces que figura una palabra o secuencia de palabras, y categorización de los datos embebidos en él de acuerdo a criterios como tema, orden alfabético, autor, frecuencia de aparición, medio de publicación, procedencia geográfica, entre otros. Otras características de un corpus son la

autenticidad de sus datos –es decir, que cuente con muestras reales del uso de la lengua en estudio para la construcción de teorías en el corpus–, textos elegidos de acuerdo a criterios lingüísticos o extralingüísticos –como archivos, colecciones o bibliotecas electrónicas– que tengan relación con el objetivo que busca la herramienta, tamaño determinado por la recopilación de textos (lo cual puede ser antes y así se crea un corpus finito, o puede mantenerse abierto, como el caso del proyecto CREA, de la Real Academia Española), y el poseer textos que respondan a parámetros estadísticos que aseguren que estos son una demostración de la variedad de la lengua de estudio –como la obra de un autor determinado, un período de tiempo, un género, etc.–, es decir, una muestra seleccionada de la lengua del corpus (Villayandre, 2008).

3.5.3 Lexicón

Es un listado de palabras de un lenguaje –un vocabulario– acompañado de un poco de conocimiento acerca de cómo cada palabra es utilizada. Puede ser general, como un lexicón de palabras comunes en inglés, o específica del dominio, como un lexicón de términos de odontología respecto a un idioma, por ejemplo. Las palabras que son de interés son, usualmente, palabras de contenido o de clase abierta, tales como sustantivos, verbos y adjetivos, en lugar de las palabras con función gramatical o de clase cerrada, como artículos, pronombres y preposiciones, cuyo comportamiento está más estrechamente ligado a la gramática de la lengua. También puede incluir expresiones de varias palabras, como verbos compuestos ("ha consumido", "hubiera estado intentando") u otras expresiones comunes ("¡Feliz Navidad!") (Hirst, 2009).

Cabe añadir que, de acuerdo a lo que reconoce Evens (como se citó en Moreno, 1998), este modelo es una pieza elemental para los investigadores enfocados en el estudio de problemas relacionados al lenguaje, ya que en el ámbito lingüístico este es propicio para un análisis completo de la sintaxis y la semántica. En el ambiente computacional, los lexicones son fundamentales para construir sistemas informáticos (sea o no que empleen o apoyen al estudio de NLP) que propicien, fortalezcan y amenicen la interacción humano-computadora; para ello, deben contener la mayor información léxica y detalle posible que corresponda y represente lo mejor posible a la realidad humana (Moreno, 1998).

Las tres principales características con las que un lexicón debe contar son: tener una sólida base psicolingüística, ser reutilizable y ser distribuible (Moreno, 1998).

3.6 Keyword Spotting Technique o Técnica de Reconocimiento de Palabras

Clave

Este método, empleada en diferentes investigaciones (Francisco y Gervás, 2006; García y Alías, 2008; Shivhare et al., 2015) tiene como finalidad hallar un conjunto de palabras clave, simples o compuestas, en la conversación de entrada, el cual puede ser oral (audio) o escrita (texto) (Heracleous y Shimizu, 2003). Si se desea orientar para hallar emociones en un texto, tal como lo hicieron Shivhare et al. (2015), por ejemplo, se definen estas palabras emocionales que deben detectarse en el texto, las cuales pueden ser verbos, sustantivos o adjetivos, para luego ser clasificadas en categorías de afecto, como alegría, tristeza, enojo, etc. Esta técnica es muy popular por su bajo costo computacional (García y Alías, 2008) y por su accesibilidad, es decir, su facilidad para integrarlo a un sistema.

Sin embargo, esta técnica posee dos grandes debilidades: su análisis se basa en las características superficiales del texto, y presenta inconvenientes al detectar el cambio de polaridad en el texto ante la presencia de algunos elementos del texto; principalmente, las negaciones. En otras palabras, no puede abstraer el significado subyacente de las oraciones, y puede calificar una oración como “me siento feliz” y “no me siento feliz” con la misma calificación, solo por haber encontrado que tienen en común la palabra “feliz” (Cambria, Schuller, Xia, y Havasi, 2013).

CAPÍTULO IV: DESARROLLO DE LA SOLUCIÓN PROPUESTA

4.1 Métodos de investigación

Ante la problemática descrita en esta investigación, se implementará un sistema capaz de detectar el estado de ánimo de una persona aplicando Sentiment Analysis (SA). De la Figura 3.5, acerca de las técnicas de SA, se han elegido dos enfoques: aquel basado en solo el concepto de NLP, y aquel cuyo análisis de texto se aplica a nivel de documento; esto quiere decir que este sistema empleará métodos para procesar el texto que deben principalmente poder trabajar con un lexicón en español. Por esta razón, se hará uso de la técnica de detección de palabras clave y la técnica de SA basada en solo NLP y en uso de lexicones. Asimismo, este sistema debe contar con una clasificación predeterminada de los estados de ánimo. Por ello, para esta investigación se definirá que el estado de ánimo tiene cuatro tipos, o estados, según el “Modelo Bidimensional de los estados de ánimo” propuesto por André (2010): tensión, sufrimiento, apatía y bienestar; adicionalmente, se añadirá a la clasificación un tipo de estado de ánimo, denominado “tranquilidad” (Figura 4.1), el cual se define por tener igual cantidad de estados de ánimo positivos y negativos (André, 2010).

Figura 4.1

Tipos de estados de ánimo reconocidos en el sistema propuesto



Nota: Tipos de estados de ánimo establecidos en el sistema propuesto, según "Modelo bidimensional de los estados de ánimo".

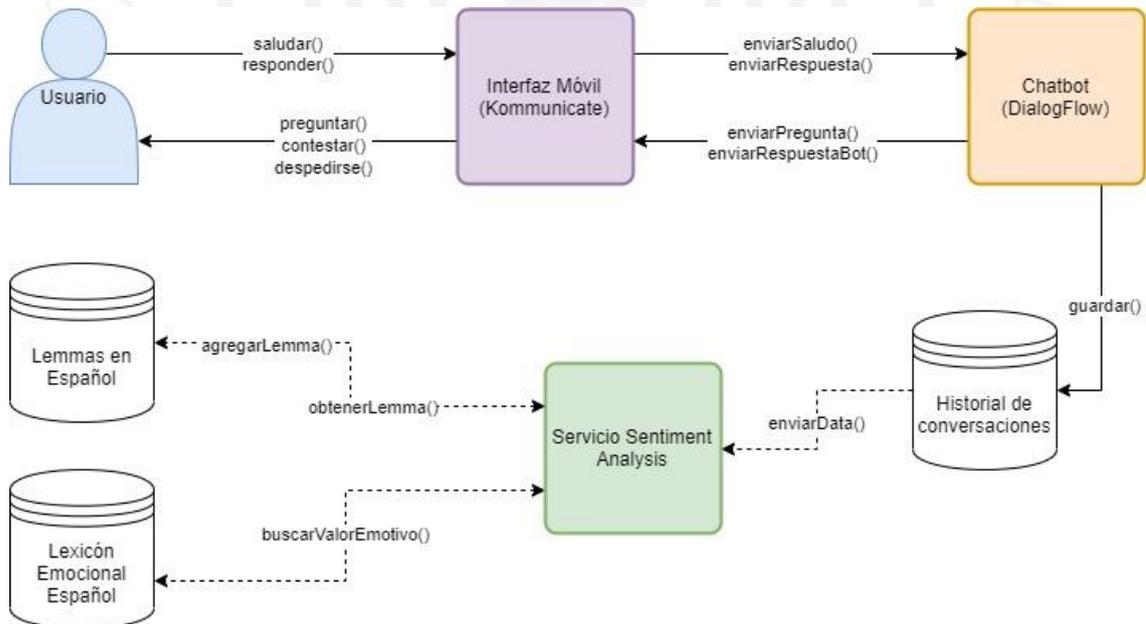
Fuente: André, C. (2010)

Elaboración propia.

Se procesarán 49 conversaciones, producto del lenguaje y de los estados de ánimo de una persona (Maturana, como se citó en Merino, 2015). Estas conversaciones se producirán entre un bot desarrollado para esta investigación, y los participantes mediante un interfaz móvil que permita la comunicación de ambas partes. Para ello, se establecerá en el bot las frases que debe emitir para saludar, preguntar por el estado de ánimo actual del usuario y terminar la conversación; asimismo, este bot recibirá las respuestas que registre el usuario para luego procesarlas por el Servicio Implementado el cual será desarrollado en base a SA con el objetivo de detectar el estado de ánimo del usuario (Figura 4.2). Con este fin, se utilizará DialogFlow (Google, 2016) para la creación y configuración del bot y la estructura de diálogo que debe seguir, y Kommunicate (Kommunicate, 2017) para la integración entre el bot y la interfaz móvil, mediante la cual el usuario ingresará su respuesta; gracias a estas dos herramientas, se logrará el nexo para que el usuario registre cómo se siente y el programa detecte el estado de ánimo del usuario.

Figura 4.2

Arquitectura del sistema propuesto



Nota: La arquitectura de la solución propuesta está conformada por el chatbot alojado en un servicio, la interfaz móvil conectada al chatbot, y el servicio de Sentiment Analysis en español, el cual interactuará con tres fuentes de datos: el Lexicón Emocional Español, la fuente de Lemmas en español, y la base de datos del propio sistema.

Elaboración propia.

Las respuestas textuales obtenidas a partir de la interacción entre el usuario y el bot serán procesados por un servicio de SA en textos, el cual será implementado en esta investigación. Se usarán las arquitecturas brindadas en el trabajo de Shirvahe et al. (2015), en la investigación de García y Alías y en el estudio de Francisco y Gervás (2006), para diseñar el flujo de trabajo de esta tesis; estas tres arquitecturas han sido previamente descritas en el Estado del Arte de la presente investigación.

El bot seguirá la siguiente estructura de conversación validada y aprobada por expertos: Saludo y pregunta inicial, Pregunta de fondo, Despedida y cierre. Las frases y preguntas que compondrán estas tres partes será de forma impersonal, para así evitar sesgo de información. En la primera parte, el bot saludará al usuario y le hará una pregunta que requiera una respuesta corta o de pocos detalles (ver Tabla 4.1).

Tabla 4.1

Frases correspondientes al saludo y pregunta inicial

Saludo e inicio de conversación. Pregunta inicial.	
Ítem	Dimensión
¡Hola! ¿Cómo estás?	Tranquilidad
¡Hola! ¿Hasta ahora cómo vas?	Tranquilidad
¡Buen día! ¿Cómo estás?	Tranquilidad
¡Hola! Hasta ahora, ¿cómo va tu día?	Tranquilidad
¡Hola! ¿Cómo te encuentras?	Tranquilidad
¡Buen día! ¿Cómo te va?	Tranquilidad
¡Hola! ¿Qué tal el día?	Tranquilidad
¡Buen día! ¿Cómo te encuentras?	Tranquilidad
¡Hola! Cuéntame, ¿cómo te ha ido?	Tranquilidad

Nota: Inicio de conversación entre usuario y bot. Está compuesto por un saludo y una pregunta inicial, la cual invita al usuario a indicar cómo se siente en ese momento.

Elaboración propia.

Una vez que se haya culminado la primera parte, el bot emitirá la Pregunta de fondo en función a la respuesta brindada por el usuario. Es decir, si la respuesta de la primera pregunta está relacionada a mayor presencia de estados de ánimo positivos, se

le escogerá aleatoriamente una pregunta del conjunto de preguntas que esté relacionada a mayor cantidad de estados de ánimo positivos (Tabla 4.2).

Tabla 4.2

Preguntas de fondo ante respuestas positivas a la pregunta inicial

Pregunta de Fondo en caso la respuesta a la pregunta inicial es positiva	
Ítem	Dimensión
¿En serio? ¡Me alegro mucho! ¿Podrías comentarme lo sucedido?	Bienestar
¡Qué bueno! Dime, ¿qué ha pasado?	Tensión
¡Qué bien! Cuéntame, ¿qué ha sucedido?	Tensión
¡Me alegro mucho por ti! Dime, ¿qué ha pasado?	Bienestar
¡Qué bueno! Descríbeme por qué es un buen día para ti	Bienestar
¿De verdad? ¡¡Qué maravilla!! Dime, ¿cómo fue?, ¿qué pasó?	Tensión

Nota: Todas las preguntas de esta sección corresponden a categorías de mayor presencia de estados de ánimo positivos. Estos son Bienestar y Tensión.

Elaboración propia.

De igual manera sucederá cuando se ingrese una respuesta con mayor presencia de estados de ánimo negativo (Tabla 4.3).

Tabla 4.3

Preguntas de fondo ante respuestas negativas a la pregunta inicial

Pregunta de Fondo en caso la respuesta a la pregunta inicial es negativa	
Ítem	Dimensión
Por favor, dime lo que ha sucedido	Depresión
¿Y por qué lo sientes así? Coméntame lo que piensas	Apatía
¿Por qué lo sientes así?	Depresión
¿Qué cosas te afectaron?	Apatía
¿Qué ha pasado?	Depresión
¿Podrías comentarme qué cosas te afectaron?	Apatía

Nota: Todas las preguntas de esta sección corresponden a categorías de mayor presencia de estados de ánimo negativos. Estos son Depresión y Apatía.

Elaboración propia.

También el bot actuará de la misma forma cuando el usuario ingrese una respuesta indiferente; en otras palabras, una respuesta que posea misma presencia de estados de ánimo positivo y negativo (Tabla 4.4).

Tabla 4.4

Preguntas de fondo ante respuestas indiferentes (ni positivas ni negativas) a la pregunta inicial

Pregunta de Fondo en caso la respuesta a la pregunta inicial es indiferente	
Ítem	Dimensión
¿Cómo te has ido sintiendo?	Tranquilidad
Coméntame, qué has estado haciendo	Tranquilidad
Dime, ¿cómo te has ido sintiendo durante el día?	Tranquilidad
Coméntame, ¿cómo te sientes?	Tranquilidad
¿Qué ha pasado?	Tranquilidad
¿Cómo te sientes?	Tranquilidad

Nota: Todas las preguntas de esta sección corresponden al estado de ánimo de Tranquilidad.
Elaboración propia.

El bot podrá emitir varias y diferentes preguntas iniciales y de fondo. Por último, el bot emitirá una despedida en tono impersonal, la cual invita al usuario a escribir en otro momento o a terminar la conversación (Tabla 4.5).

Tabla 4.5

Frases correspondientes a la despedida y cierre de conversación

Despedida y cierre de conversación	
Ítem	Dimensión
Si deseas, puedes escribirme cómo te va. Buen día.	Bienestar
Muchas gracias por tus respuestas.	Tranquilidad
Gracias. Que tengas un buen día.	Bienestar
Gracias por responder. Buen día.	Tranquilidad
Puedes comentarme cómo te va cuando desees.	Bienestar
Muchas gracias, y buen día.	Tranquilidad

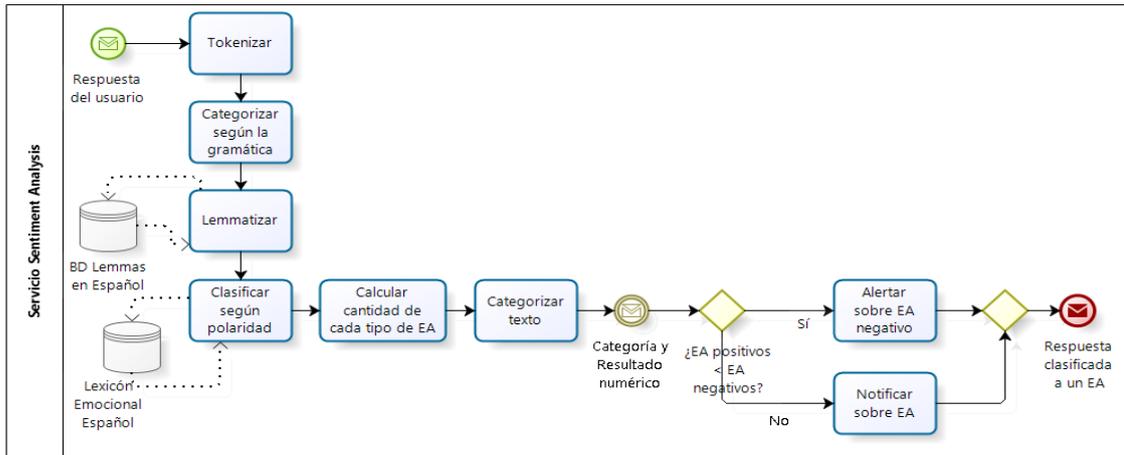
Nota: Frases dichas por el bot para invitarlo a escribir en otro momento, despedirse y culminar la conversación en ese instante.

Elaboración propia.

El flujo de trabajo será el siguiente: el texto del usuario será primero dividido en palabras ('tokens') para luego asociar cada palabra a su etiqueta gramatical; seguido, se identificará cuál es el lemma correspondiente a cada palabra según su categoría gramatical, para después, con ella, contabilizar cuántos lemmas buscados en el Lexicón Emocional Español corresponden a estados de ánimo positivos y cuántos negativos. Finalmente, en función a la cantidad de estados de ánimo según su polaridad y a la diferencia de cantidades entre ellos, se categoriza el tipo de estado de ánimo del texto y se le notifica al usuario (Figura 4.3).

Figura 4.3

Mapa de procesos del servicio del Servicio de Sentiment Analysis



Nota: Diagrama de procesos del Servicio de Sentiment Analysis en textos en español. En función a la metodología elegida, categorizará el texto de acuerdo al estado de ánimo que corresponde. Elaboración propia.

El proceso de detección del estado de ánimo comenzará con el ingreso de la respuesta textual del usuario al sistema; inmediatamente, aplicando la técnica Keyword Spotting Technique –la cual ha sido descrita en el Marco Teórico de este trabajo, en la sección del mismo nombre–, la respuesta textual será tokenizada en palabras y se categorizará cada una según la gramática española; en otras palabras, la respuesta textual será dividida en palabras y, de cada una de ellas, se determinará si es un sustantivo, un adjetivo, un signo de puntuación, etc. (ver Tabla 4.6). Para lograrlo, se utilizará la librería Stanford CoreNLP (Manning, Surdeanu, Bauer, Finkel, Bethard, y McClosky, 2014; Stanford NLP Group, 2014).

Tabla 4.6

Categorización gramatical de palabras en idioma español con Stanford CoreNLP

Tag	Description	Example(s)
Adjectives		
ao0000	Adjective (ordinal)	<i>primera, segundo, últimos</i>
aq0000	Adjective (descriptive)	<i>populares, elegido, emocionada, andaluz</i>
Conjunctions		
cc	Conjunction (coordinating)	<i>y, o, pero</i>
cs	Conjunction (subordinating)	<i>que, como, mientras</i>
Determiners		
da0000	Article (definite)	<i>el, la, los, las</i>
dd0000	Demonstrative	<i>este, esta, esos</i>
de0000	"Exclamative" (TODO)	<i>qué (¡Qué pobre!)</i>
di0000	Article (indefinite)	<i>un, muchos, todos, otros</i>
dn0000	Numeral	<i>tres, doscientas</i>

(continúa)

(continuación)

do0000	Numeral (ordinal)	<i>el 65 aniversario</i>
dp0000	Possessive	<i>sus, mi</i>
dt0000	Interrogative	<i>cuántos, qué, cuál</i>
Punctuation		
f0	Other	<i>y, @</i>
faa	Inverted exclamation mark	<i>¡</i>
fat	Exclamation mark	<i>!</i>
fc	Comma	<i>,</i>
fca	Left bracket	<i>[</i>
fct	Right bracket	<i>]</i>
fd	Colon	<i>:</i>
fe	Double quote	<i>"</i>
fg	Hyphen	<i>-</i>
fh	Forward slash	<i>/</i>
fia	Inverted question mark	<i>¿</i>
fit	Question mark	<i>?</i>
fp	Period / full-stop	<i>.</i>
fpa	Left parenthesis	<i>(</i>
fpt	Right parenthesis	<i>)</i>
fra	Left guillemet / angle quote	<i>«</i>
frc	Right guillemet / angle quote	<i>»</i>
fs	Ellipsis	<i>..., etcétera</i>
ft	Percent sign	<i>%</i>
fx	Semicolon	<i>;</i>
fz	Single quote	<i>'</i>
Interjections		
i	Interjection	<i>ay, ojalá, hola</i>
Nouns		
nc0000	Unknown common noun (neologism, loanword)	<i>minidisc, hooligans, re-flotamiento</i>
nc0n000	Common noun (invariant number)	<i>hipótesis, campus, golf</i>
nc0p000	Common noun (plural)	<i>años, elecciones</i>
nc0s000	Common noun (singular)	<i>lista, hotel, partido</i>
np00000	Proper noun	<i>Málaga, Parlamento, UFINSA</i>
Pronouns		
p0000000	Impersonal <i>se</i>	<i>Se</i>
pd000000	Demonstrative pronoun	<i>éste, eso, aquellas</i>
pe000000	"Exclamative" pronoun	<i>Qué</i>
pi000000	Indefinite pronoun	<i>muchos, uno, tanto, nadie</i>
pn000000	Numeral pronoun	<i>dos miles, ambos</i>
pp000000	Personal pronoun	<i>ellos, lo, la, nos</i>
pr000000	Relative pronoun	<i>que, quien, donde, cuales</i>
pt000000	Interrogative pronoun	<i>cómo, cuánto, qué</i>
px000000	Possessive pronoun	<i>tuyo, nuestra</i>
Adverbs		
rg	Adverb (general)	<i>siempre, más, personalmente</i>
rn	Adverb (negating)	<i>No</i>
Prepositions		
sp000	Preposition	<i>en, de, entre</i>
Verbs		
va00000	Verb (unknown)	<i>Should</i>
vag0000	Verb (auxiliary, gerund)	<i>Habiendo</i>
vaic000	Verb (auxiliary, indicative, conditional)	<i>habría, habríamos</i>
vaif000	Verb (auxiliary, indicative, future)	<i>habrá, habremos</i>
vaii000	Verb (auxiliary, indicative, imperfect)	<i>había, habíamos</i>
vaip000	Verb (auxiliary, indicative, present)	<i>ha, hemos</i>
vais000	Verb (auxiliary, indicative, preterite)	<i>hubo, hubimos</i>

(continúa)

(continuación)

vam0000	Verb (auxiliary, imperative)	<i>Haya</i>
van0000	Verb (auxiliary, infinitive)	<i>Haber</i>
vap0000	Verb (auxiliary, participle)	<i>Habido</i>
vasi0000	Verb (auxiliary, subjunctive, imperfect)	<i>hubiera, hubiéramos, hubiese</i>
vasp0000	Verb (auxiliary, subjunctive, present)	<i>haya, hayamos</i>
vmg0000	Verb (main, gerund)	<i>dando, trabajando</i>
vmic0000	Verb (main, indicative, conditional)	<i>daría, trabajaríamos</i>
vmif0000	Verb (main, indicative, future)	<i>dará, trabajaremos</i>
vmii0000	Verb (main, indicative, imperfect)	<i>daba, trabajábamos</i>
vmip0000	Verb (main, indicative, present)	<i>da, trabajamos</i>
vmis0000	Verb (main, indicative, preterite)	<i>dio, trabajamos</i>
vmm0000	Verb (main, imperative)	<i>da, dé, trabaja, trabajos, trabajemos</i>
vmn0000	Verb (main, infinitive)	<i>dar, trabajar</i>
vmp0000	Verb (main, participle)	<i>dado, trabajado</i>
vmsi0000	Verb (main, subjunctive, imperfect)	<i>diera, diese, trabajáramos, trabajésemos</i>
vmsp0000	Verb (main, subjunctive, present)	<i>dé, trabajemos</i>
vsg0000	Verb (semiauxiliary, gerund)	<i>Siendo</i>
vsic0000	Verb (semiauxiliary, indicative, conditional)	<i>sería, serían</i>
vsif0000	Verb (semiauxiliary, indicative, future)	<i>será, seremos</i>
vsii0000	Verb (semiauxiliary, indicative, imperfect)	<i>era, éramos</i>
vsip0000	Verb (semiauxiliary, indicative, present)	<i>es, son</i>
vsis0000	Verb (semiauxiliary, indicative, preterite)	<i>fue, fuiste</i>
vsm0000	Verb (semiauxiliary, imperative)	<i>sea, sé</i>
vsn0000	Verb (semiauxiliary, infinitive)	<i>Ser</i>
vsp0000	Verb (semiauxiliary, participle)	<i>Sido</i>
vssf0000	Verb (semiauxiliary, subjunctive, future)	<i>Fuere</i>
vssi0000	Verb (semiauxiliary, subjunctive, imperfect)	<i>fuera, fuese, fuéramos</i>
vssp0000	Verb (semiauxiliary, subjunctive, present)	<i>sea, seamos</i>
Dates		
w	Date	<i>octubre, jueves, 2002</i>
Numerals		
z0	Numeral	<i>547.000, 04, 52, 52</i>
zm	Numeral qualifier (currency)	<i>dólares, euros</i>
zu	Numeral qualifier (other units)	<i>km, cc</i>
Other		
word	Emoticon or other symbol	<i>:), ®</i>

Nota: Códigos asociados a las categorías gramaticales establecidos en la librería Stanford CoreNLP para textos en idioma español.

Fuente: Stanford NLP Group (2014). Recuperado de <https://nlp.stanford.edu/software/spanish-faq.html>

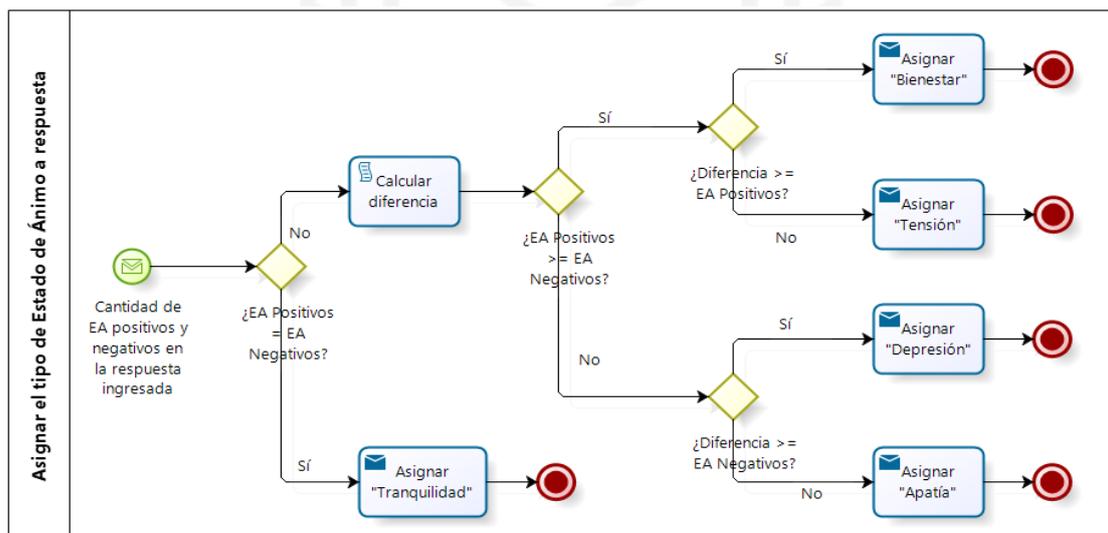
Una vez que haya sido tokenizada y categorizada gramaticalmente, cada palabra de la respuesta ingresada será lematizada: a comparación del stemming – procedimiento que consiste en retirar los sufijos y prefijos de la palabra para conseguir el morfema o raíz–, se reducirá cada palabra al nivel de su morfema en base a un previo análisis morfológico (Risueño, 2018). Para conseguir la lematización de cada palabra, se empleará en esta sección la base de datos abierta –en inglés, Open Database– de nombre “lemmatization-es”; esta base de datos fue creada por Michal Měchura (2018), contiene solo lemmas en español y está asociado a 565 292 palabras.

Después de haber sido lematizada, se contabilizará cuántas palabras positivas y negativas contiene la respuesta ingresada por el usuario. En este módulo, se empleará el Lexicón Emocional para discernir si la palabra que se está analizando está relacionada al estado de ánimo positivo, o al estado de ánimo negativo, o no está relacionada a ninguno de los dos. El Lexicón Emocional que se utilizará en esta investigación pertenece a la empresa Autoritas Consulting (Rangel, 2016), y, a comparación del lexicón elaborado por el equipo del investigador Javier Redondo en 2007 y conformado por 1034 palabras (Redondo, Fraga, Padrón, y Comesaña, 2007), este lexicón está compuesto por 1300 palabras positivas y 1270 palabras negativas (Autoritas Consulting, 2015). Cabe mencionar que el contenido del Lexicón Emocional será previamente lematizado, antes que comience todo el proceso de detección de estados de ánimo, para poder analizar la respuesta ingresada ya lematizada.

Finalmente, ya calculadas las cantidades de palabras positivas y negativas existentes en la respuesta ingresada y la diferencia entre estos valores numéricos, se categorizará cuál es el estado de ánimo presente en la respuesta ingresada. Para ello, sobre la base del modelo propuesto por André (2010), para el sistema propuesto se aplicará el flujo condicional diagramado en la Figura 4.4.

Figura 4.4

Flujo condicional para la categorización del estado de ánimo



Nota: Flujo condicional para determinar el estado de ánimo. Basado en el Modelo Bidimensional de André (2010).

Fuente: André, C. (2010).

Elaboración propia.

4.2 Alcance

- La solución no dará ningún diagnóstico psicológico ni clínico. No se busca ni pretende sustituir al especialista. Tampoco brindará recomendaciones de actividades o acciones que afecten su estado de ánimo.
- La solución cumplirá solo con las funciones de registro y detección del estado de ánimo de una persona.
- El sistema solo procesará respuestas textuales en idioma español, ingresadas por usuarios de habla hispana.
- El sistema hará el cálculo y detección de estado de ánimo de solo lo ingresado al sistema.

4.3 Supuestos

- El usuario ingresa respuestas honestas y detalladas al sistema.
- Domina el idioma español.

4.4 Riesgos

El primer riesgo es el registro de respuestas deshonestas por parte del usuario. Ante ello, se plantea en esta investigación el supuesto que el usuario hace uso del sistema de forma honesta, es decir, indica cómo realmente se siente en ese instante. Asimismo, se establece el alcance que se procesará solo las palabras que el usuario ingrese al sistema.

Otro riesgo es la presencia de jergas y expresiones producidas por la mezcla con otras lenguas –como el Spanglish– en el habla cotidiana. Para ello, el sistema las tratará como si fuesen una palabra más: determinará cuál es su categoría gramatical, luego hallará sus lemmas en la Base de datos de Lemmas en español y por último obtendrá sus polaridades buscándolo en el Lexicón Emotivo, en el caso que primero encuentre esas palabras allí.

CAPÍTULO V: PRUEBAS Y RESULTADOS

Para el desarrollo y validación de la investigación, los instrumentos que se utilizaron fueron el Servicio Implementado en base a Sentiment Analysis y Keyword Spotting Technique para el procesamiento de textos en español y detección de los estados de ánimos presente, y la aplicación móvil para la recopilación de conversaciones dadas entre una persona y el bot, el cual siguió el flujo de diálogo e hizo las preguntas establecidas y validadas por un experto.

En cuanto a los participantes, estos fueron 49 hombres y mujeres hispanohablantes entre 23 y 33 años de edad que laboraban en una empresa privada del rubro tecnológico de Lima Metropolitana. Esta muestra, así como su tamaño, también fue validada por el experto. En líneas generales, estas personas, por medio de la aplicación móvil implementada, respondieron qué hicieron durante el día y cómo se sintieron en el momento que se les aplicó la prueba. Previamente, se les informó que su participación era voluntaria y anónima; asimismo, se les indicó que su respuesta no debía contener jergas ni emoticones ni errores ortográficos, era libre de juicio –en otras palabras, no existía respuesta correcta o incorrecta–, y debía estar escrita en idioma español. En total, se obtuvieron 49 conversaciones, producto de la interacción entre cada usuario con el bot mediante la aplicación móvil.

Como se mencionó en el capítulo anterior, este bot siguió el siguiente flujo de conversación: primero, saludar y preguntar cómo se siente; luego, preguntar por qué se siente así y qué actividades realizó durante el día; finalmente, da una frase de despedida. En todas las conversaciones, el bot inició el diálogo y ha estado enviando preguntas en función a la respuesta que ha brindado el usuario (Figura 5.1).

Figura 5.1

Modelo de conversación entre un participante y el bot



Nota: Conversación entre usuario y bot a través de la aplicación empleada, siguiendo la estructura de conversación instaurada. Las preguntas que emite el bot dependen de las palabras que use el usuario para indicar cómo se siente. Esta conversación pasa al historial de conversaciones para luego ser procesada en el Servicio Implementado, el cual se encarga de determinar el estado de ánimo del participante. Elaboración propia.

Junto con el experto, se corroboró que las preguntas sean de carácter impersonal y que no presenten palabras que sesgue información; es decir, que oriente al participante a un tipo de respuesta. Se utilizó el Test de Personalidad basado en la tipología de Jung y el indicador de Myers-Briggs (HumanMetrics Inc., 1998) como referencia para la elaboración de las preguntas; el experto se encargó de revisar si las preguntas redactadas correspondían, en primer lugar, a la definición de estados de ánimo según André (2010) y, en segundo lugar, si correspondía a uno de los cinco tipos de estados de ánimo que define el mismo autor (ver Anexo 4).

Luego de haber recopilado las conversaciones, solo las respuestas de los usuarios ante las preguntas que realizó el bot se ingresaron al sistema de detección de estados de ánimo, el cual siguió el proceso explicado previamente. Se revisó cada respuesta y se corrigió faltas ortográficas y uso incorrecto de signos de puntuación por parte del usuario para que el sistema pueda clasificar adecuadamente los estados de ánimo; un ejemplo de estas faltas ortográficas es la siguiente: “un.baile”.

Previamente a la aplicación del sistema propuesto, se realizó una prueba piloto del Servicio Implementado con la participación de 30 personas con las mismas características de la muestra, las cuales realizaron una conversación con el bot configurado por medio de la aplicación móvil desarrollada. Este piloto se realizó con el fin de validar los instrumentos desarrollados, verificar si las preguntas que conforman el diálogo del bot ayudan a que los usuarios brinden data no sesgada e identificar mejoras del sistema propuesto.

Una vez recopilado las conversaciones de los participantes de la prueba piloto, se validó los resultados obtenidos bajo la supervisión de un experto. Para ello, se entrevistó a 10 usuarios que participaron en el piloto; se les preguntó cuál era su estado de ánimo el día que realizaron la prueba, cómo se sentían y con cuál estado de ánimo de la teoría de André (2010) sintieron que estaba asociado su estado de ánimo en ese momento. Al contrastar los resultados del sistema propuesto con los de la entrevista, se comprobó que el Servicio Implementado clasificó correctamente en un 70% de los casos. En la Tabla 5.1 se muestra la validación del piloto ejecutado.

Tabla 5.1

Comparación entre los resultados de la entrevista y los resultados del Servicio Implementado

N° usuario	Resultado de Entrevista	Resultado de Servicio
Usuario 1	Tensión	Tensión
Usuario 2	Tensión	Apatía
Usuario 3	Depresión	Depresión
Usuario 4	Apatía	Apatía
Usuario 5	Tranquilidad	Bienestar
Usuario 6	Bienestar	Bienestar
Usuario 7	Bienestar	Tensión
Usuario 8	Bienestar	Bienestar
Usuario 9	Bienestar	Bienestar
Usuario 10	Apatía	Apatía

Nota: Cuadro comparativo entre los resultados obtenidos de la entrevista realizada a los participantes de la prueba piloto y los resultados obtenidos del procesamiento del Servicio Implementado. Elaboración propia.

Con base en los resultados obtenidos de la validación y de los datos plasmados en la Tabla 5.1, se armó la siguiente matriz de confusión, en la cual se muestran cuántas conversaciones que categorizó el Servicio Implementado coincidieron con las respuestas de los participantes, y cuántos no coincidieron (Tabla 5.2). Se calculó que la precisión –en inglés, “accuracy” – del sistema propuesto ante estos 10 casos es de 0.7; es decir, tal como se mencionó previamente en este capítulo, el Servicio Implementado acertó en el 70% de los casos presentados en la prueba piloto.

Tabla 5.2

Matriz de confusión de la validación de la prueba piloto

		Predicción				
		Tranquilidad	Apatía	Depresión	Tensión	Bienestar
Actual	Tranquilidad	0	0	0	0	1
	Apatía	0	2	0	0	0
	Depresión	0	0	1	0	0
	Tensión	0	1	0	1	0
	Bienestar	0	0	0	1	3

Nota: Matriz de confusión de la validación de los resultados de la prueba piloto, tanto de la predicción que realizó el Servicio Implementado como de las entrevistas realizadas a los participantes. Elaboración propia.

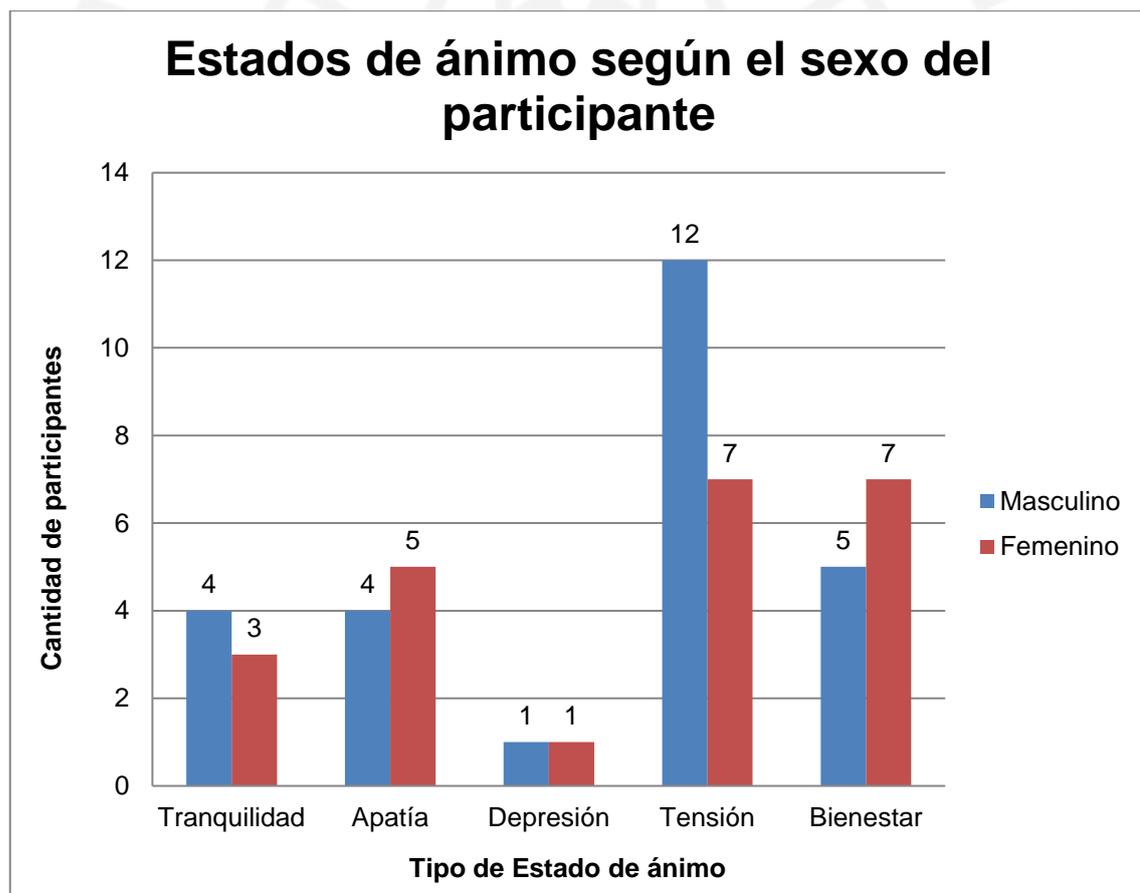
Ya demostrada la validez de los instrumentos, se procedió con la aplicación del sistema propuesto en la muestra de 49 hombres y mujeres hispanohablantes de entre 23

y 33 años de edad que laboran en una empresa privada del rubro tecnológico de Lima Metropolitana.

A partir de la clasificación de la información obtenida según el sexo de la muestra, se pudo comprobar que los participantes de sexo masculino presentan, en mayor proporción, el estado de ánimo de tensión, mientras que los participantes de sexo femenino poseen estados de ánimo de tensión y de bienestar. Esto demuestra que los participantes, en el momento en que se les aplicó la prueba, poseían gran cantidad de estados de ánimo positivos, puesto que en los estados de ánimo de Tensión y Bienestar existe una numerosa cantidad de estados de ánimo positivos, tal como lo indica André (2010) en su Modelo Bidimensional de los Estados de Ánimo (Figura 5.2).

Figura 5.2

Estados de ánimo detectados según el sexo del participante



Nota: Gráfico de barras relacionado a los estados de ánimo presentes en el personal según su sexo. En los participantes de sexo masculino, predomina la Tensión; mientras que en los participantes de sexo femenino, los dos estados de ánimo que lideran son Tensión y Bienestar. Elaboración propia.

También se pudo observar los estados de ánimo que se presentaron con mayor frecuencia en los participantes durante las fechas que se les aplicó la prueba. Como se puede apreciar, el estado de ánimo de Tensión predomina sobre la muestra. En segundo lugar está Bienestar, y seguido, los estados de ánimo de Apatía, Tranquilidad y Depresión, respectivamente (Figura 5.3). Este es un indicativo que los participantes, durante ese momento, presentaron porcentajes elevados de estados de ánimo positivo y negativo, siendo el estado de ánimo positivo el que sobresalió.

Figura 5.3

Proporción de la presencia de los tipos de estados de ánimo en los participantes



Nota: Proporción de los tipos de estados de ánimo presente en los participantes. Se puede apreciar que la mayoría de los participantes poseía un estado de ánimo de Tensión el día que realizó la prueba. Elaboración propia.

Ante este indicativo, se contabilizó y rankeó las 10 palabras que más se utilizaron y estuvieron relacionadas tanto a estados de ánimo positivos como estados de ánimo negativos, según la base de datos léxica emocional proporcionada por Autoritas

Consulting (Autoritas Consulting, 2015). Se puede observar que las palabras que corresponden a estados de ánimo positivos en el lexicón que se usaron recurrentemente fueron “bien”, “gracia” y “sentir” (Figura 5.4).

Figura 5.4

Palabras del lexicón emocional asociadas a estados de ánimo positivos más utilizadas



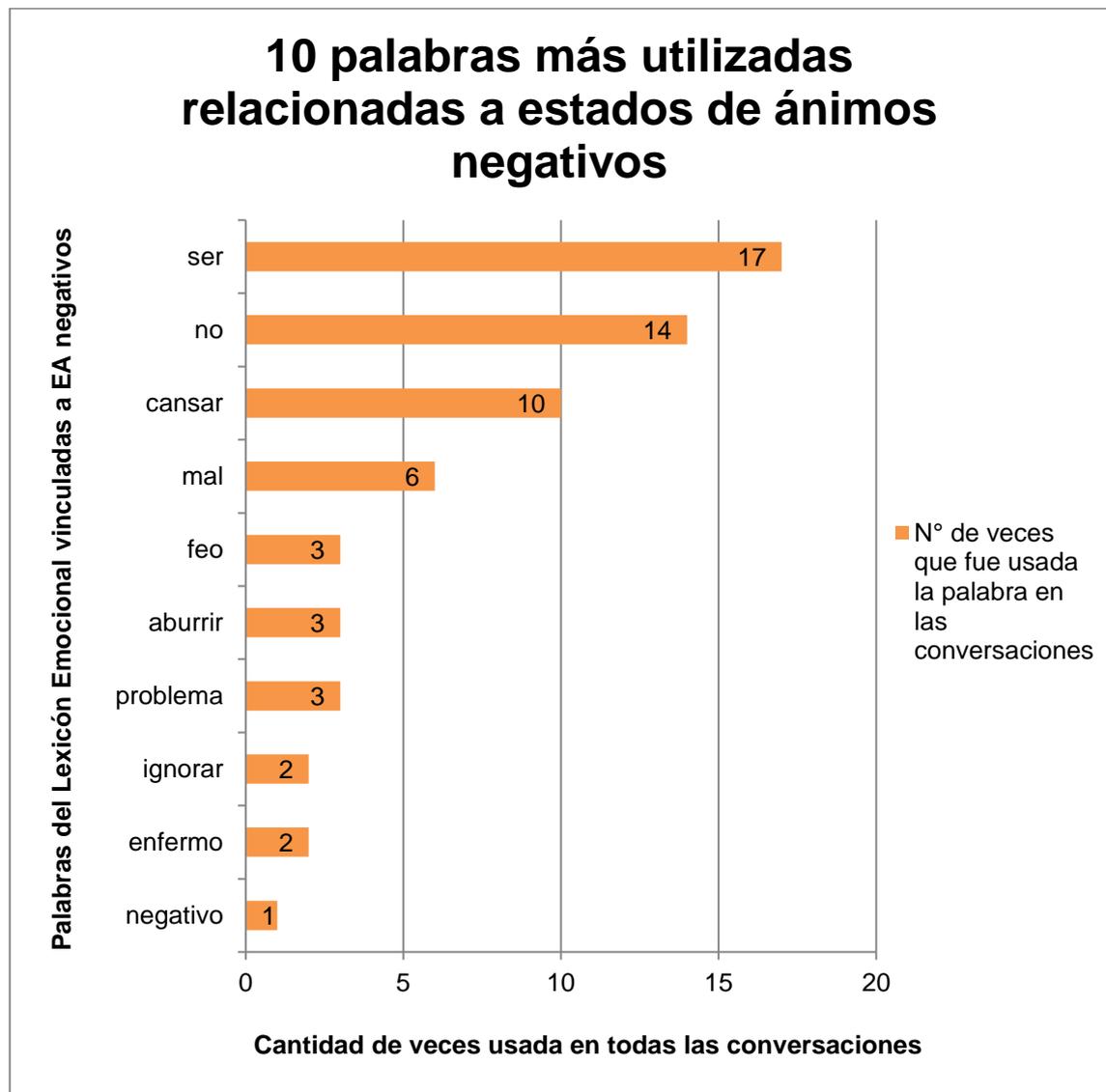
Nota: Clasificación de las 10 palabras más empleadas en las conversaciones y que estén asociadas a estados de ánimo positivos en el lexicón emocional.

Elaboración propia.

En cambio, la palabra “ser” es la palabra que corresponde a estados de ánimo negativos en el lexicón más empleada, seguido de “no” y “cansar”. Esto se debe a la relación directa entre los lemmas detectados en las conversaciones y los lemmas correspondientes a los elementos del lexicón, pues la palabra “ser” es tanto un sustantivo común singular como un verbo irregular infinitivo (Figura 5.5).

Figura 5.5

Palabras del léxico emocional asociadas a estados de ánimo negativos más utilizadas



Nota: Clasificación de las 10 palabras más empleadas en las conversaciones y que estén asociadas a estados de ánimo negativos en el léxico emocional.
Elaboración propia.

Finalmente, se pudo comprobar que parte del tiempo de ejecución del Servicio Implementado es logarítmico, debido a que se realizan entre dos y tres búsquedas binarias. La primera es la búsqueda de lemma en el léxico de lemmas: primero se busca en el archivo que contiene todos los lemmas de la categoría gramatical que le asoció la librería Stanford CoreNLP (Manning et al., 2014; Stanford NLP Group, 2014); de no encontrar el lemma en ese archivo, realiza la búsqueda en el repositorio de todos los lemmas. La segunda búsqueda binaria es la búsqueda del lemma identificado en el léxico emocional, cuyo contenido pasó previamente por un proceso de

lemmatización. Se validó el costo computacional que implica usar lexicones y la técnica de encontrar palabras clave (KST) al comparar la aplicación de búsqueda binaria con la aplicación de búsqueda lineal, y se comprobó que, al emplear búsqueda lineal, el sistema tarda 20 segundos en procesar un texto de 200 palabras, aproximadamente, mientras que, al aplicar búsqueda binaria, se minimiza el tiempo: el sistema tarda 5 segundos en procesar un texto de 200 palabras.



CONCLUSIONES

- El uso de tecnologías y metodologías relacionadas a Inteligencia Artificial, como Sentiment Analysis y NLP, favorecen a estudios y disciplinas vinculadas al comportamiento humano. A partir de esta investigación, se da la posibilidad de generar soluciones y productos para empresas de habla española con relación al estado de ánimo de las personas de su entorno. Por ejemplo, se puede desarrollar para el área de Recursos Humanos una aplicación que les permita conocer el estado de ánimo de sus colaboradores para detectar puntos críticos y desarrollar estrategias para mejorar el clima laboral, o para realizar campañas de prevención ante problemas de estado de ánimo.
- El modelo de este sistema sirve también para el desarrollo de aplicaciones que favorezcan al autoconocimiento de una persona. A partir de la prueba piloto y la prueba final de la aplicación, se pudo apreciar que el producto de la interacción de cada usuario con el bot, mediante las preguntas que este le hacía, para cada persona resultó ser un reflejo de sí mismo en cuanto a su estado de ánimo: su forma de escribir ya mostraba cómo se encontraba anímicamente.
- La técnica de Keyword Spotting Technique está estrechamente relacionada con la búsqueda binaria, y aplicarla para la búsqueda de cada palabra del texto en tratamiento en las diferentes bases de datos léxicas favorece a la agilidad de análisis y procesamiento de textos. Usar esta técnica junto con las metodologías de Sentiment Analysis relacionadas a solo los conceptos de NLP y el uso de bases de datos léxicas generan menor costo computacional y favorecen al análisis de textos a nivel sintáctico.
- Elegir conversaciones para esta investigación muestra que detectar el estado de ánimo ya no solo se puede realizar en textos estáticos, de carácter impersonal y ajenos a la cotidianidad, como son las publicaciones en blogs, artículos y cuentos. No solo se ha logrado procesar textos estrechamente relacionados al quehacer de cada persona, a sus costumbres y

a su manera de percibir la realidad: es posible detectar cómo se siente en ese instante de tiempo.

- Se ha demostrado que existen librerías que favorecen el tratamiento, análisis y procesamiento de textos en español.
- Se ha demostrado que existen bases de datos léxicas en español, como la base de datos de lemmas y la base de datos léxica emocional, y también se ha demostrado que existen librerías y herramientas capaces de tratar textos en idioma español, como es el caso de Stanford CoreNLP.



RECOMENDACIONES

- Para las siguientes etapas de implementación, si la orientación de este sistema va para el autoconocimiento, es conveniente que el resultado sea enviado durante la conversación; asimismo, también se busca que el aplicativo realice acciones una vez que ya detecte el estado de ánimo de su usuario; por ejemplo, puede enviar un listado de actividades que le sugiera a su usuario ante el estado de ánimo que presenta en ese instante.
- Para optimizar la precisión, se recomienda que el análisis realizado por el servicio implementado para detectar estados de ánimo en texto efectúe análisis a nivel semántico.
- Se sugiere profundizar el estudio de procesamiento del lenguaje natural en el idioma español, y generar productos que favorezcan a las investigaciones que estén enfocadas a este idioma.
- Se pueden mejorar o producir nuevas soluciones capaces de tratar textos en idioma español. Por ejemplo, una base de datos léxica emocional con más de 5000 palabras, o una base de datos léxica de jergas y neologismos de una comunidad de habla hispana. Incluso, se puede investigar cómo hacer que estas bases de datos estén en constante actualización, pues el lenguaje natural cambia en función a sus hablantes; de igual manera, estas bases de datos deben adecuarse al cambio del día a día.

GLOSARIO DE TÉRMINOS

Bot.- Aféresis de robot. Software que emula un comportamiento humano para establecer comunicación con un usuario con el fin de brindarle información y ayuda en las aplicaciones donde funciona (Haj-Saleh, 2017).

Chatbot. - Tecnología que permite establecer una conversación entre una persona y un sistema informático con el fin de solicitar información o de realizar una acción (Rodríguez, 2017).

Lematización. - Proceso textual donde las palabras correspondientes a un mismo paradigma flexivo o derivativo son llevadas a una forma estándar que representa a toda la clase; su finalidad es hallar la raíz morfológica de las palabras (Bassi, 1999).

Lemma. - Es la palabra utilizada como entrada en los diccionarios de lengua: el infinitivo para las conjugaciones verbales, el masculino singular para adjetivos, etc. (Bassi, 1999).

Lingüística. - Chomsky (2003) lo describe como la ciencia empírica que cuestiona la naturaleza del lenguaje humano como fenómeno natural que está representado de forma concreta (mediante una formalización determinada) en el cerebro (Chomsky, citado en Ramírez, 2013).

Stemming. - Conocido también como flexionador (García y Alías, 2008), es una variación de la lematización; este proceso consiste en encontrar el lexema de las palabras mediante la reducción de sufijos y prefijos presentes (Bassi, 1999).

Valencia. - Es la atracción (valencia positiva) o aversión (valencia negativa) que provoca un determinado objeto o situación. También se emplea para clasificar emociones y, en caso exista conflicto entre valencias positivas y negativas, se denomina “ambivalencia” (Pérez Porto y Gardey, 2012).

REFERENCIAS

- André, C. (2010). *Los estados de ánimo: El aprendizaje de la serenidad*. Barcelona: Editorial Kairós.
- AR Productions Inc. (4 de Abril de 2018). *Registro de Estados de Ánimo*. Recuperado el 15 de Diciembre de 2018, de Google Play:
<https://play.google.com/store/apps/details?id=arproductions.andrew.moodlog>
- Aspyre Solutions. (3 de Agosto de 2010). *My Mood Tracker*. Recuperado el 15 de Diciembre de 2018, de Aspyre Apps: <http://aspyreapps.com/projects/my-mood-tracker>
- Autoritas Consulting. (2015). Recuperado el 25 de Setiembre de 2016, de Autoritas:
<http://www.autoritas.net/>
- Ávila, J. (20 de Abril de 2016). *¿Qué es la ontología en informática?* Recuperado el 19 de Enero de 2019, de Aplicaciones Informáticas:
<https://basicinfoweb.wordpress.com/2016/04/20/que-es-la-ontologia-en-informatica/>
- Bassi, A. (1999). *Lematización basada en análisis no supervisado de corpus*. Santiago de Chile: Departamento de Ciencias de la Computación - Universidad de Chile.
- Beenish, C. (04 de Agosto de 2016). Daylio: mood-quantification for a less stressful you. *mHealth*, 2.
- Blasco, L. (6 de Noviembre de 2018). ¿Cuántas lenguas originarias existen en Perú y cuáles de éstas se hablan en otros países? *BBC News*.
- Brackett, M., y Stern, R. (2017). *Mood Meter: Build Emotional Intelligence to Last a Lifetime*. Recuperado el 14 de Diciembre de 2018, de Mood Meter:
<http://moodmeterapp.com/>
- Bradley, M., y Lang, P. (1999). *Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings*. Florida: Universidad de Florida.
- Brafman, S. (27 de Octubre de 2017). *STIGMA APP: Help To Improve Your Mental Health*. Recuperado el 14 de Diciembre de 2018, de Mental Health Center:
<http://www.mentalhealthcenter.org/stigma-app-mental-health-html/>
- Buenaga, M. d. (Enero de 2001). Integración de técnicas de procesamiento del lenguaje natural para la recuperación de información en bibliotecas de componentes software (Tesis Doctoral). 206. Madrid, España: Universidad Complutense de Madrid.
- Caldeira, C., Chen, Y., Chan, L., Pham, V., Chen, Y., y Zheng, K. (2017). Mobile apps for mood tracking: an analysis of features and user reviews. *AMIA Annual Symposium Proceedings* (págs. 495–504). Irvine: American Medical Informatics Association.

- Cambria, E. (2016). Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems*, 31(2), 102-107.
- Cambria, E., Schuller, B., Xia, Y., y Havasi, C. (21 de Febrero de 2013). New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems*, XXVIII(2), 15-21.
- Carbajales, S. (23 de Noviembre de 2018). *Es complicado que las máquinas nos entiendan con lenguaje natural, pero si es en español, aún más*. Obtenido de Xataka.
- Catalán, S. (10 de Julio de 2017). *Cómo influye el estado de ánimo en nuestra vida*. Recuperado el 1 de Marzo de 2019, de Cipsia Psicólogos: <https://www.cipsiapsicologos.com/blog/como-influye-el-estado-de-animo-en-nuestra-vida/>
- Collomb, A., Costea, C., Joyeux, D., Hasan, O., y Brunie, L. (2014). *A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation*. Centre National de la Recherche Scientifique, Laboratoire d'InfoRmatique en Image et Systèmes d'information. Lyon: Rapport de recherche RR-LIRIS-2014-002.
- Corbin, J. A. (1 de Setiembre de 2017). *Los 8 tipos de emociones (clasificación y descripción)*. Recuperado el 3 de Julio de 2018, de Psicología y Mente: <https://psicologiymente.com/psicologia/tipos-de-emociones>
- cppreference.com . (4 de Mayo de 2012). *Tiempo de ejecución*. Recuperado el 19 de Noviembre de 2019, de cppreference.com : <https://es.cppreference.com/w/cpp/complexity>
- Darlington, K. (04 de Enero de 2017). *El comienzo de la era de la Inteligencia Artificial*. Recuperado el 17 de Junio de 2019, de BBVA OpenMind: <https://www.bbvaopenmind.com/tecnologia/inteligencia-artificial/el-comienzo-de-la-era-de-la-inteligencia-artificial/>
- Desmet, P. (2015). Design for Mood: Twenty Activity-Based Opportunities to Design for Mood Regulation. *International Journal of Design*, 1-19.
- Díaz-Agudo, B., Gervás, P., Hervás, R., y Peinado, F. (2004). Story Plot Generation based on CBR. En A. Macintosh, R. Ellis, y T. Allen (Ed.), *12th Conference on Applications and Innovations in Intelligent Systems*, (págs. 1-14). Cambridge.
- Domenach, F., Charmarai, P., Savva, A., y Christou, C. (2015). Felt ??? A social feeling app. *International Conference on Interactive Mobile Communication, Technologies and Learning (IMCL)*, (págs. 163-166). Thessaloniki.
- Fernández, F. d. (2011). *Integración de métodos para la desambiguación del sentido de las palabras en el contexto del procesamiento del lenguaje natural*. La Habana: D - Instituto Superior Politécnico José Antonio Echeverría (CUJAE).
- Francisco, V., y Gervás, P. (Setiembre de 2006). Análisis de dependencias para la marcación de cuentos con emociones. *Procesamiento del Lenguaje Natural*, 137-144.

- Francisco, V., Hervás, R., y Gervás, P. (Setiembre de 2005). Expresión de emociones en la síntesis de voz en contextos narrativos. *Simposio de Computación Ubicua e Inteligencia Ambiental*. Madrid.
- Gallup. (2018). *2018 Global Emotions Report*. Gallup Press.
- García, D., y Alías, F. (Abril de 2008). Identificación de emociones a partir de texto usando desambiguación semántica. (S. E. Natural, Ed.) *SEPLN*(40), 75-81.
- Goasduff, L. (22 de Enero de 2018). *Update: Emotion AI will personalize interactions*. Recuperado el 17 de Enero de 2019, de Smarter with Gartner: <https://www.gartner.com/smarterwithgartner/emotion-ai-will-personalize-interactions/>
- Goleman, D. (2003). *Emociones destructivas: Un diálogo científico con el Dalai Lama*. (D. González Raga, y F. Mora, Trans.) Nueva York: Bantam Books.
- Good to Hear. (2012). Obtenido de Good to Hear. Improving life through technology: <http://goodtohear.co.uk/happiness>
- Google. (2016). *Overview: DialogFlow*. Recuperado el Marzo de 2019, de DialogFlow: <https://dialogflow.com/>
- Haj-Saleh, A. (2017). Qué son exactamente los 'bots' y cómo funcionan. *Revista GQ*.
- Heracleous, P., y Shimizu, T. (2003). An efficient keyword spotting technique using a complementary language for filler models training. *Conference: 8th European Conference on Speech Communication and Technology, EUROSPEECH 2003 - INTERSPEECH 2003*, (págs. 921-924). Geneva.
- Hirst, G. (2009). Ontology and the Lexicon. En S. Staab, y R. Studer, *Handbook on Ontologies* (Segunda ed., págs. 269-292). Toronto: Springer.
- HumanMetrics Inc. (1998). *Test de Tipología de Jung*™. Recuperado el 3de Setiembre de 2019, de HumanMetrics: <http://www.humanmetrics.com/personalidad>
- Jhangiani, R., y Tarry, H. (2014). The Role of Affect: Moods and Emotions. En C. Stangor, *Principles of Social Psychology – 1st International Edition* (Primera ed.).
- Karimkhan, P. (5 de Octubre de 2017). *What is the relation between sentiment analysis, natural language processing and machine learning? [Publicación en un foro online]*. Recuperado el 18 de Enero de 2019, de Quora: <https://www.quora.com/What-is-the-relation-between-sentiment-analysis-natural-language-processing-and-machine-learning>
- Kilô, E. (1 de Junio de 2017). Los altibajos del estado de ánimo. *Vanguardia*, pág. 1.
- Kommunicate. (2017). *Home: Kommunicate*. Recuperado el 30 de Abril de 2019, de Kommunicate: <https://www.kommunicate.io/>

- Kumar, E. (2011). *Natural Language Processing*. Nueva Delhi: I K International Publishing House.
- Larter, R. (1 de Febrero de 2012). Toolmaker Talk: Ross Larter (MoodPanda). (R. Mehta, Entrevistador)
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Toronto, Canadá: Synthesis Lectures on Human Language Technologies.
- Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., y McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, (págs. 55-60). Baltimore.
- Měchura, M. (11 de Mayo de 2018). lemmatization-lists. *Lemmatization Lists*.
- Michael, A. (17 de Marzo de 2018). *What is the relation between sentiment analysis, natural language processing and machine learning? [Publicación en un foro online]*. Recuperado el 18 de Enero de 2019, de Quora: <https://www.quora.com/What-is-the-relation-between-sentiment-analysis-natural-language-processing-and-machine-learning>
- Moreno, A. (1998). El lexicón en la lexicografía computacional: adquisición y representación de la información léxica. *Alfinge. Revista de Filología*, 249-272.
- Mulcrone, K. (2012). Detecting emotion in text. University of Minnesota–Morris CS Senior Seminar Paper.
- Munezero, M., Suero Montero, C., Sutinen, E., y Pajunen, J. (2014). Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text. *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING*, 5(2), 101-111.
- Oliver, A., Moré, J., y Climent, S. (2008). *Traducción y tecnologías*. Barcelona: UOC (UNIVERSITAT OBERTA DE CATALUNYA).
- Oracle Corporation. (1996). *Java*. Recuperado el 19 de Junio de 2016, de Java: <https://www.java.com/es/>
- Pérez Porto, J., y Gardey, A. (2012). *Definición de Valencia*. Recuperado el Julio de 2016, de Definición.DE: <http://definicion.de/valencia/>
- Pérez, M. Á., y Redondo, M. (2006). Procesos de valoración y emoción: características, desarrollo, clasificación y estado actual. *Revista Electrónica de Motivación y Emoción*, 1-47.
- Polanco Fernández, D. F. (2000). Capítulo 2. Lenguajes Naturales y Lenguajes Formales. En D. F. Polanco Fernández, *Evaluación y mejora de un sistema automático de análisis sintagmático* (págs. 11-27). Madrid, España: Universidad Politécnica de Madrid.
- Porter, M., MacFarlane, A., y Boulton, R. (Febrero de 2002). *Snowball Stem*. Recuperado el 10 de Setiembre de 2016, de Snowball: <http://snowballstem.org/>

- Ramírez González, B. (Diciembre de 2013). Hacia un modelo computacional unificado del lenguaje natural. *Linguamática*, V(2), 91-100.
- Rangel, F. (Junio de 2016). *Author Profiling en Social Media: Identificación de Edad, Sexo y Variedad del Lenguaje*. Tesis doctoral, Universitat Politècnica de València, Valencia.
- Redondo, J., Fraga, I., Padrón, I., y Comesaña, M. (2007). The Spanish adaptation of ANEW (Affective Norms for English Words). *Behavior Research Methods*, 6.
- Reeve, J. (2010). *Motivación y emoción* (Quinta ed.). México D.F: McGraw-Hill.
- Rincón, S. (13 de Abril de 2013). 6 causas de los cambios de humor. *SinEmbargo MX*.
- Risueño, T. (28 de Febrero de 2018). *What is the difference between stemming and lemmatization?* Recuperado el 1 de Marzo de 2019, de Bitext: <https://blog.bitext.com/what-is-the-difference-between-stemming-and-lemmatization/>
- Rodríguez, Á. (2017). *Chatbots: lo que debes saber para incluirlos en tu estrategia de contenidos*. Recuperado el 10 de Diciembre de 2019, de 40defiebre: <https://www.40defiebre.com/chatbots-estrategia-contenidos>
- Rovira, I. (Abril de 2018). *Cambios de humor: qué son, y 9 consejos para controlarlos*. Recuperado el 10 de Marzo de 2019, de Psicología y Mente: <https://psicologiymente.com/psicologia/cambios-de-humor>
- Santamaría, I. (2007). *El argot y las jergas*. Madrid: Liceus.
- Schimelpfening, N. (28 de Noviembre de 2018). *Possible causes of Mood Swings*. Recuperado el 2 de Marzo de 2019, de VeryWell Mind: <https://www.verywellmind.com/what-are-mood-swings-1067178>
- Shivhare, S. N., Garg, S., y Mishra, A. (2015). EmotionFinder: Detecting Emotion From Blogs and Textual Documents. *International Conference on Computing, Communication and Automation (ICCCA2015)*, (págs. 52-57). Noida.
- Silver, N. (16 de Febrero de 2018). *What Can Cause Rapid Mood Swings?* Recuperado el 2 de Abril de 2019, de Healthline: <https://www.healthline.com/health/rapid-mood-swings>
- Stanford NLP Group. (2014). *Spanish FAQ for Stanford CoreNLP, parser, POS tagger, and NER*. Recuperado el Octubre de 2016, de The Stanford Natural Language Processing Group: <https://nlp.stanford.edu/software/spanish-faq.html>
- Thayer, R. (1998). *El origen de los estados de ánimo cotidianos*. Barcelona, España: Paidós.
- The Apache Software Foundation. (Noviembre de 2010). *Apache Jena*. Recuperado el 19 de Junio de 2016, de Apache Jena: <https://jena.apache.org/>
- Villayandre, M. (2008). Lingüística con corpus. *E.H. Filología*(30), 329-349.

Windwer, M. (2017). *Moodtrack Social Diary*. Recuperado el 14 de Diciembre de 2018, de Google Play:
<https://play.google.com/store/apps/details?id=com.moodtrak.diaryyhl=en>

Wong, K. (19 de Octubre de 2015). *Track Your Happiness Helps You Understand What Makes You Happy*. Recuperado el Junio de 2016, de LifeHacker:
<http://lifehacker.com/track-your-happiness-helps-you-understand-what-makes-yo-1737038850>



BIBLIOGRAFÍA

- "*Quantified Self on Happiness*". (25 de Julio de 2011). Recuperado el Junio de 2016, de The Ultimate Answer: <http://www.theultimateanswer.org/quantified-self-on-happiness.html>
- AffectSense. (s.f.). *Singularity University: AffectSense*. Recuperado el Junio de 2016, de <http://singularityu.org/companies/affect-sense/>
- Álvarez, E., Almenar, C., y Regalado, P. (2015). *Envejecimiento Saludable. Estado de ánimo y depresión en las personas mayores*. Barcelona: Fundació Salut i Envelliment UAB.
- Ball, J. (2006). *Can NLP Systems be a Cognitive Black Box?* Arizona.
- Basis Technology Corp. (2016). Recuperado el 29 de Agosto de 2016, de Rosette Text Analytics: <https://www.rosette.com/>
- cppreference.com . (4 de Mayo de 2012). *Tiempo de ejecución*. Recuperado el 19 de Noviembre de 2019, de [cppreference.com](https://es.cppreference.com/w/cpp/complexity) :
- El Tiempo. (15 de Octubre de 2018). El MIT creará una universidad dedicada a la inteligencia artificial. *El Tiempo*.
- Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press.
- Fessl, A., Rivera-Pelayo, V., Pammer, V., y Braun, S. (2012). Mood Tracking in Virtual Meetings. *Proceedings of the 7th European conference on Technology Enhanced Learning*, (págs. 377-382). Karlsruhe.
- Mehta, R. (9 de Mayo de 2012). Toolmaker Talk. (M. Forrest, Entrevistador)
- Mitkova, A. (2007). El léxico juvenil por áreas temáticas. *Tonos Digital, Revista Electrónica de Estudios Filológicos*, 1-11.
- Nesse, R. (24 de Noviembre de 1991). What is Mood for? *Psychology*.
- Otlogetswe, T. J., y Ramaeba, G. (2014). Developing a Campus Slang Dictionary for the University of Botswana. *Lexikos*(1), 350-361.
- Puente, L. (12 de Enero de 2015). *El lenguaje de los millenials*. Obtenido de ProExpansión: <http://proexpansion.com/es/articles/1005-el-lenguaje-de-los-millenials>
- Purdy, M., Zealley, J., y Maseli, O. (18 de Noviembre de 2019). *The Risks of Using AI to Interpret Human Emotions*. Recuperado el 29 de Noviembre de 2019, de Harvard Business Review: <https://hbr.org/2019/11/the-risks-of-using-ai-to-interpret-human-emotions>

- Ramírez Vásquez, N. M. (2009). Aspectos semánticos en la jerga estudiantil universitaria, sede de Guanacaste. *Káñina*, 177-184.
- Rodríguez Ponce, M. I. (30 de Enero de 2012). Apreciaciones sobre elementos valorativos y usos fácticos en el estilo comunicativo juvenil. *Sintagma*(24), 8-21.
- Saggion, H., y Funk, A. (2010). Interpreting SentiWordNet for Opinion Classification. *Proceedings of the International Conference on Language Resources and Evaluation*, (págs. 1129 - 1133). Barcelona.
- Sánchez, H. (16 de Noviembre de 2016). *¿Qué es el Sentiment Analysis?* Recuperado el 5 de Marzo de 2019, de PostedIn: <https://www.postedin.com/blog/que-es-el-sentiment-analysis/>
- Sandstrom, G., Lathia, N., Mascolo, C., y Rentfrow, P. (2017). Putting mood in context: Using smartphones to examine how people feel in different locations. *Journal of Research in Personality*, 96-101.
- Santamaría González, F. (27 de Setiembre de 2012). *La autocuantificación: el emergente y cuantificado en los procesos vitales*. Recuperado el 10 de Agosto de 2018, de Blog de Fernando Santamaría: <http://fernandosantamaria.com/blog/tag/auto-cuantificado/>
- Sarzotti, F. (2018 de Enero de 2018). Self-Monitoring of Emotions and Mood Using a Tangible Approach. *Computers*, 1-28.



ANEXOS

ANEXO 1: FORMATO DE CONSENTIMIENTO INFORMADO

UNIVERSIDAD DE LIMA
FACULTAD DE INGENIERÍA Y ARQUITECTURA
CARRERA DE INGENIERÍA DE SISTEMAS

CONSENTIMIENTO INFORMADO

Este trabajo de investigación corresponde a la estudiante **Stephany Alessandra Morzán Fuentes** de la Universidad de Lima, de la carrera de Ingeniería de Sistemas identificada con DNI Nro. **70308494**, quien está asesorada por el docente **Juan Manuel Gutiérrez Cárdenas** con DNI Nro. **29515539**.

Yo, _____ DNI: _____

ACEPTO participar en el proceso de validación del trabajo de investigación titulado “**Detección de Estados de Ánimo mediante Sentiment Analysis en hispanohablantes**”, que se encuentra en evaluación experimental, mediante un prototipo de **modelo computacional**. Comprendo que mi participación es totalmente libre, anónima y voluntaria; y que aún después de iniciada la investigación, puedo decidir suspender mi participación en cualquier momento, sin necesidad de expresar razón y sin que eso ocasione algún perjuicio.

Los responsables del proyecto podrán divulgar la información que se genere producto de mi participación en la investigación, mas no podrán divulgar mi información personal.

Declaro que mi participación no implica ninguna retribución, por tratarse de una investigación académica.

Nombre y apellido:

ANEXO 2: URLS DEL REPOSITORIO DE CONVERSACIONES Y DEL SERVICIO IMPLEMENTADO

Repositorio de historial de conversaciones:

https://github.com/stephmorzan/laTitu/blob/master/historial_conversaciones.txt

Repositorio de código fuente del Servicio Implementado:

<https://github.com/stephmorzan/laTitu>



**ANEXO 3: RESULTADOS DE LAS CONVERSACIONES
TRATADAS POR EL SISTEMA PROPUESTO**

N°	Sexo de Usuario	Cant. Palabras Positivas	Cant. Palabras Negativas	Palabras Positivas	Palabras Negativas	Resultados
1	M	0	1	-	cansar	Apatía
2	M	1	1	bien	cansar	Tranquilidad
3	F	2	0	bien, trabajar	-	Tensión
4	M	2	1	bueno, bien	no	Tensión
5	F	3	1	bien, gracia, trabajar	ser	Tensión
6	M	6	0	bien, bien, tranquilo, ayudar, calor, gracia	-	Bienestar
7	M	4	7	sentir, bien, sentir, sentir	mal, ser, no, ser, no, ser, ser	Apatía
8	F	4	2	bien, gracia, bien, bueno, novio	triste	Bienestar

(continúa)

(continuación)

9	M	2	1	bien, bien	problema	Tensión
10	F	1	2	gracia	no, no	Apatía
11	F	3	0	bien, trabajar, gracia	-	Bienestar
12	F	0	0	-	-	Tranquilidad
13	F	1	3	bien	mal, no, mal	Apatía
14	F	2	0	bien, bien	-	Tensión
15	M	2	0	bien, jugar	-	Tensión
16	F	3	1	bien, interesante, avanzar	no	Tensión
17	M	1	0	calor	-	Tensión
18	M	0	3	-	aburrir, aburrir, cansar	Apatía
19	M	0	2	-	matar, venganza	Apatía

(continúa)

(continuación)

20	M	2	0	trabajar, bien	-	Tensión
21	M	2	0	trabajar, bien	-	Tensión
22	M	2	2	bien, querer	cansar, ser	Tranquilidad
23	F	5	1	bien, curar, bueno, bien decir	enfermedad	Bienestar
24	M	2	2	trabajar, jugar	enfermo, enfermo	Tranquilidad
25	M	1	5	bueno	ser, responsabilid ad, no, faltar, malograr	Depresión
26	M	1	0	bien	-	Tensión
27	M	3	1	bien, tranquilo, práctico	ser	Tensión
28	M	6	2	tranquilo, mejor, bien, gusto, querer, bueno	ser, ser	Bienestar
29	M	3	2	bien, bien, tranquilo	no, ser	Tensión
30	M	3	0	bien, trabajar, probar	-	Bienestar

(continúa)

(continuación)

31	F	2	2	bien, productivo	ser, pesar	Tranquilidad
32	M	6	1	bien, gracia, bueno, memorable, disfrutar, diversión	cansar	Bienestar
33	F	0	3	-	ignorar, feo, privar	Apatía
34	F	5	1	bien, gracia, calor, bien, querer	doler	Bienestar
35	F	1	1	bien	cansar	Tranquilidad
36	M	2	0	bien, divertir	-	Tensión
37	F	2	4	sentir, sentir	agotar, mal, alterar, problema	Apatía
38	F	4	1	bien, tranquilo, tranquilo, tranquilo	ser	Bienestar
39	F	5	0	bien, sentir, bien, bien, gracia	-	Bienestar
40	M	4	2	bien, sentir, optimista, apoyar	feo, conflicto	Tensión

(continúa)

(continuación)

41	F	5	1	bien, productivo, avanzar, bien, emocionar	ser	Bienestar
42	M	6	3	bien, sentir, feliz, sentir, saludable, fuerte	satisfacer, problema, no	Bienestar
43	M	5	5	dedicar, sentir, fluir, considerar, positivo	aburrir, no, estrés, ser, negativo	Tranquilidad
44	F	1	7	bien	cansar, no, mal, discriminar, feo, ignorar, mal	Depresión
45	F	2	1	tranquilo, eficiente	ser	Tensión
46	F	2	0	bien, avanzar	-	Tensión
47	M	4	2	gracia, verdad, positivo, gracia	ser, cansar	Tensión
48	F	2	3	bien, bien	cansar, cansar, no	Apatía
49	F	2	1	bien, excelente	no	Tensión

ANEXO 4: VALIDACIÓN DE LA MATRIZ DE DIÁLOGO DEL BOT PARA CRITERIO DEL EXPERTO

Matriz de diálogo de bot para Criterio de Expertos

Estimado(a) colaborador(a):

Le escribo para solicitar su apoyo como juez(a) en el proceso de validar las siguientes preguntas, las cuales han sido formuladas para que el evaluado brindara respuestas abiertas, mediante el criterio de jueces. Las preguntas están dirigidas a una población de hombres y mujeres en el rango de edad entre 23 y 33 años que trabajan en una empresa privada de rubro tecnológico en Lima Metropolitana.

Estas preguntas buscan sonsacar el estado de ánimo del evaluado en ese momento, así como saber los motivos o causas el evaluado se siente de esa manera. El Estado de Ánimo se puede definir como el estado psicológico que se ve influenciado y va de la mano de un conglomerado de sensaciones autopercibidas que determinan el bienestar psicológico, la calidad de las relaciones interpersonales y qué tan eficaz puede uno comportarse. En base a teorías evolutivas e investigaciones neurocientíficas, existen dos tipos de estados de ánimo: positivos y negativos; no obstante, estos. El comportamiento del estado de ánimo en el ser humano es bidimensional: si bien existen los estados de ánimo positivo y negativo, estos no se contrarrestan entre sí. Puede una persona tener mayor cantidad de estados de ánimo positivos que de negativos, y viceversa; asimismo, una persona puede tener la misma cantidad de estados de ánimos positivos y negativos, teniendo así un estado de ánimo “neutro”. Estos tipos de estados de ánimo, los cuales ahora los denominaré “dimensiones”, son:

1. Tensión: Existe una presencia abundante de estados de ánimo positivos, y una presencia abundante de estados de ánimo negativos. Mayor es la presencia de estados de ánimo positivos.
2. Bienestar: Existe una presencia abundante de estados de ánimo positivos, y una presencia escasa de estados de ánimo negativos.

3. Apatía: Existe una presencia escasa de estados de ánimo positivos, y una presencia escasa de estados de ánimo negativos. Mayor es la presencia de estados de ánimo negativos.
4. Depresión: Existe una presencia abundante de estados de ánimo negativos, y una presencia escasa de estados de ánimo positivos.
5. Tranquilidad: La presencia de estados de ánimo positivo y negativo es igual.

Modelo bidimensional de los estados de ánimo



En la tabla adjunta, se le presentarán las preguntas agrupadas según el flujo de conversación establecido con el evaluado. Por favor calificarlos en base a los siguientes criterios:

- Representativo al constructo: las preguntas pertenecen al constructo de Estado de Ánimo. Se solicita calificar este criterio mencionando “Sí”, si es que la

pregunta es representativo del constructo; y “No”, si es que el ítem no es representativo del constructo.

- Adecuada redacción: la formulación de las preguntas es gramaticalmente correcta y tiene un vocabulario adecuado. Deberá calificar este criterio con un “Sí”, si es que la pregunta es entendible y claro para la población dirigida, además que la pregunta le invite al usuario a ingresar una respuesta abierta, de mínimo 10 palabras; y “No”, si es que la pregunta no lo es.

En ambos casos en los que se califique “NO”, o en el caso de tener alguna observación adicional, se solicita indicar la razón o duda en la casilla de “Observaciones”.

Muchas gracias por su tiempo y valioso apoyo.

Saludos cordiales,
Stephany Morzán

Saludo e inicio de conversación. Pregunta inicial

Ítem	Dimensión	Representativo al constructo		Observaciones	Adecuada redacción		Observaciones
		Sí	No		Sí	No	
¡Hola! ¿Cómo estás?	Tranquilidad	Sí	No		Sí	No	
¡Hola! ¿Hasta ahora cómo vas?	Tranquilidad	Sí	No		Sí	No	
¡Buen día! ¿Cómo estás?	Tranquilidad	Sí	No		Sí	No	
¡Hola! Hasta ahora, ¿cómo va tu día?	Tranquilidad	Sí	No		Sí	No	
¡Hola! ¿Cómo te encuentras?	Tranquilidad	Sí	No		Sí	No	
¡Buen día! ¿Cómo te va?	Tranquilidad	Sí	No		Sí	No	
¡Hola! ¿Qué tal el día?	Tranquilidad	Sí	No		Sí	No	
¡Buen día! ¿Cómo te encuentras?	Tranquilidad	Sí	No		Sí	No	
¡Hola! Cuéntame, ¿cómo te ha ido?	Tranquilidad	Sí	No		Sí	No	

Si la respuesta a una de las preguntas de la sección de Saludo es Afirmativa:

Ítem	Dimensión	Representativo del Constructo		Observaciones	Adecuada redacción		Observaciones
		Sí	No		Sí	No	
¿En serio? ¡Me alegre mucho! ¿Podrías comentarme lo sucedido?	Bienestar	Sí	No		Sí	No	
¡Qué bueno! Dime, ¿qué ha pasado?	Tensión	Sí	No		Sí	No	
¡Qué bien! Cuéntame, ¿qué ha sucedido?	Tensión	Sí	No		Sí	No	
¡Me alegre mucho por ti! Dime, ¿qué ha pasado?	Bienestar	Sí	No		Sí	No	
¡Qué bueno! Descríbeme por qué es un buen día para ti	Bienestar	Sí	No		Sí	No	
¿De verdad? ¡¡Qué maravilla!! Dime, ¿cómo fue?, ¿qué pasó?	Tensión	Sí	No		Sí	No	

Si la respuesta a una de las preguntas de la sección de Saludo es Negativa:

Ítem	Dimensión	Representativo del Constructo		Observaciones	Adecuada redacción		Observaciones
		Sí	No		Sí	No	
Por favor, dime lo que ha sucedido	Depresión	Sí	No		Sí	No	
¿Y por qué lo sientes así? Coméntame lo que piensas	Apatía	Sí	No		Sí	No	
¿Por qué lo sientes así?	Depresión	Sí	No		Sí	No	
¿Qué cosas te afectaron?	Apatía	Sí	No		Sí	No	
¿Qué ha pasado?	Depresión	Sí	No		Sí	No	
¿Podrías comentarme qué cosas te afectaron?	Apatía	Sí	No		Sí	No	

Si la respuesta a una de las preguntas de la sección de Saludo es Indiferente (no es Afirmativa ni Negativa):

Ítem	Dimensión	Representativo del Constructo		Observaciones	Adecuada redacción		Observaciones
		Sí	No		Sí	No	
¿Cómo te has ido sintiendo?	Tranquilidad	Sí	No		Sí	No	
Coméntame, qué has estado haciendo	Tranquilidad	Sí	No		Sí	No	
Dime, ¿cómo te has ido sintiendo durante el día?	Tranquilidad	Sí	No		Sí	No	
Coméntame, ¿cómo te sientes?	Tranquilidad	Sí	No		Sí	No	
¿Qué ha pasado?	Tranquilidad	Sí	No		Sí	No	
¿Cómo te sientes?	Tranquilidad	Sí	No		Sí	No	

Despedida y cierre

Ítem	Dimensión	Representativo del Constructo		Observaciones	Adecuada redacción		Observaciones
		Sí	No		Sí	No	
Si deseas, puedes escribirme cómo te va. Buen día.	Bienestar	Sí	No		Sí	No	
Muchas gracias por tus respuestas.	Tranquilidad	Sí	No		Sí	No	
Gracias. Que tengas un buen día.	Bienestar	Sí	No		Sí	No	
Gracias por responder. Buen día.	Tranquilidad	Sí	No		Sí	No	
Puedes comentarme cómo te va cuando desees.	Bienestar	Sí	No		Sí	No	
Muchas gracias, y buen día.	Tranquilidad	Sí	No		Sí	No	