



UNIVERSIDAD DE JAÉN
Escuela Politécnica Superior (Jaén)

Trabajo Fin de Máster

ASISTENTE VIRTUAL INTELIGENTE CON RECONOCIMIENTO DE EMOCIONES

Alumno/a: Valiente Martín, Javier

Tutor/a: Prof. D. Miguel Ángel García Cumbreras
D^a. Flor Miriam Plaza del Arco

Dpto.: Informática

Mayo, 2020



Universidad de Jaén

Escuela Politécnica Superior de Jaén

Departamento de Informática

Don Miguel Ángel García Cumbreiras y Doña Flor Miriam Plaza del Arco , tutores del Proyecto Fin de Máster titulado: **“Asistente virtual inteligente con reconocimiento de emociones”**, que presenta Javier Valiete Martín, autoriza su presentación para defensa y evaluación en la Escuela Politécnica Superior de Jaén.

Jaén, Mayo de 2020

El alumno:

Los tutores:

Fdo.: Javier Valiente Martín

Fdo.: Prof. D. Miguel Ángel García Cumbreiras

Fdo.: D^a. Flor Miriam Plaza del Arco

Índice

Capítulo 1: Introducción.....	5
1.1. Introducción	5
1.2. Motivación.....	6
1.3. Propósito del proyecto	8
1.4. Objetivos del proyecto.....	9
1.5. Descripción del problema.....	9
1.6. Resultados esperados	10
1.7. Planificación.....	10
1.7.1. Estimación temporal	11
1.7.2. Estimación de costes.....	13
1.7.2.1. Costes hardware	13
1.7.2.2. Costes software	13
1.7.2.3. Costes de personal	14
1.7.2.4. Otros costes.....	15
1.7.2.5. Coste total.....	15
1.8. Estructura del proyecto	16
Capítulo 2. Asistentes virtuales inteligentes.....	18
2.1 Asistentes virtuales inteligentes y sistemas conversacionales	18
2.2. Principales asistentes virtuales inteligentes	19
2.2.1. Siri.....	19
2.2.2. Google Assistant	19
2.2.3. Cortana	20
2.2.4. Bixby	20
2.2.5. Alexa	20
2.3. Comparativa asistentes virtuales inteligentes.....	25
Capítulo 3. Reconocimiento de emociones.....	28
3.1. Modelos de emociones	29
3.1.1. Modelos de emociones discretas.....	29
3.1.2. Modelos de emociones dimensionales	31
3.2. Estudio de técnicas para el reconocimiento de emociones	33
3.2.1. Basado en lexicón	33
3.2.2. Basado en Machine Learning	34
3.2.3. Enfoque híbrido	34

3.3. Estudio y análisis de lexicones.....	35
3.3.1. EmoSentNet	35
3.3.2. LIWC	36
3.3.3. NRC Emotion lexicon (Emolex)	36
3.3.4. SEL	38
Capítulo 4. Desarrollo del proyecto.....	40
4.1. Metodología para el desarrollo de software y método a seguir.....	40
4.1.1. Metodología.....	40
4.1.2. Método ágil.....	42
4.2. Historias de usuario	45
4.3. Propuesta de solución	48
4.4. Descripción de la solución	48
4.4.1. Iteración 1	49
4.4.2. Iteración 2	51
4.4.3. Iteración 3	53
4.4.4. Iteración 4	55
4.4.5. Iteración 5	57
4.5. Herramientas para la implementación del sistema	59
4.5.1. Alexa Skills Kit.....	59
4.5.2. Python.....	60
Capítulo 5. Pruebas de validación y usabilidad.	62
5.1. Validación	62
5.2. Usabilidad	62
Capítulo 6. Conclusiones y trabajos futuros.....	67
6.1. Conclusiones	67
6.2. Trabajos futuros	68
Bibliografía	69
Anexo A. Manual de instalación	71
Anexo B. Manual de usuario	73
Anexo C. Índice de ilustraciones	78
Anexo D. Índice de tablas.....	80

Capítulo 1: Introducción

En este capítulo se presentan los principales aspectos que han dado origen a la elaboración del proyecto **“Asistente virtual inteligente con reconocimiento de emociones”**.

Este trabajo trata de utilizar dos de las áreas de la informática con más auge en la actualidad, los Asistentes Virtuales Inteligentes (AVI) y el Procesamiento del Lenguaje Natural (PLN). Con el desarrollo de este trabajo se han adquirido una serie de conocimientos de ambas áreas no aprendidas durante el máster. Por un lado, me he iniciado en la programación de utilidades para los AVI y por otro he mejorado y adquirido nuevas habilidades del PLN. Ambas áreas como ya se ha dicho tienen un enorme auge actualmente, y si se siguen desarrollando como hasta ahora, tendrán unos resultados y nuevas utilidades muy importantes para diversas necesidades de la sociedad.

1.1. Introducción

La tecnología avanza a pasos agigantados en la actualidad, una de las áreas con más desarrollo es la Inteligencia Artificial (IA).

Dentro de la IA una de las tareas con mayor proyección de crecimiento es la programación para AVI. Estos dispositivos, activados en su mayoría por voz, están aumentando en variedad, visibilidad de cara al público y principalmente en número.

El desarrollo de la IA junto con el PLN está ayudando también al desarrollo de estos asistentes, avanzando en el reconocimiento de la voz, la bidireccionalidad y un aprendizaje constante, haciendo que la interacción con el humano sea más acertada con el paso del tiempo y consiguiendo que cada vez se vea a las máquinas como alguien con quien poder entablar una conversación, o preguntar cualquier cosa como si de un humano se tratase.

Por otro lado, como ya sabemos, el desarrollo de la Web 1.0 hacia la Web 2.0 ha provocado que la cantidad de contenido que se alberga en Internet ya sea nuevo o compartido, este creciendo de forma exponencial en los últimos años. Millones de personas expresan sus sentimientos a diario, ya sea en sus redes sociales, en

conversaciones, artículos formales y un largo sinfín de contenido, el cual queda almacenado en la web, formando parte del denominado Big Data.

De todo este conocimiento es posible extraer multitud de información acerca de quien lo escribe, lo que quiere expresar y como lo expresa, pero no se trata de una tarea fácil, ya que los humanos no son capaces de procesar un volumen tan extenso de información. Sin embargo, las computadoras sí tienen esa gran capacidad de procesamiento de información, por lo que son utilizadas para aprovechar parte de dicho contenido, en concreto los sentimientos y emociones que quiere transmitir el usuario. Para ello, existe la tarea de clasificación de emociones, que se ayuda de técnicas de PLN y de aprendizaje automático para obtener dicha emoción. Esta tarea pertenece a un área mucho más amplia del PLN, el análisis de sentimientos, que puede tener diferentes aplicaciones como obtener información sobre la satisfacción del cliente sobre un producto, ayudar a la hora de elegir una canción, recomendar productos e incluso predecir trastornos de salud mental. Sin embargo, llevar a cabo esta tarea no es fácil para una computadora, ya que no es capaz de entender el lenguaje humano. Es por ello que cada día se enfatice más en que las computadoras sean capaces de procesar el sentido de un texto.

1.2. Motivación

Uno de los objetivos del PLN y por tanto de la tarea de clasificación de emociones, es dar una utilidad a la ingente cantidad de información que podemos encontrar en Internet actualmente, la cual puede ayudar de manera considerable en diferentes campos.

Por otro lado, en la actualidad son cada vez más las personas que utilizan los AVI de manera cotidiana para todo tipo de tareas, como guardar un recordatorio, poner música, realizar una llamada, etc. Y no solo son muchos los asistentes utilizados hoy en día si no que se espera que el crecimiento sea aún mayor para los próximos años¹, como se muestra en la siguiente ilustración (Ilustración 1.1). Este hecho lleva consigo un rápido crecimiento del número de utilidades capaces de ofrecernos, abriendo nuevos caminos a la hora de utilizarlos. La principal motivación de este proyecto consiste en crear una utilidad para uno de estos AVI, en concreto para el Amazon

¹ <https://es.statista.com/estadisticas/972995/asistentes-virtuales-en-uso-en-el-mundo/>

Alexa que sea capaz de unir las ventajas y usos de esta herramienta junto con las que proporciona el PLN mediante el reconocimiento y clasificación de las emociones.

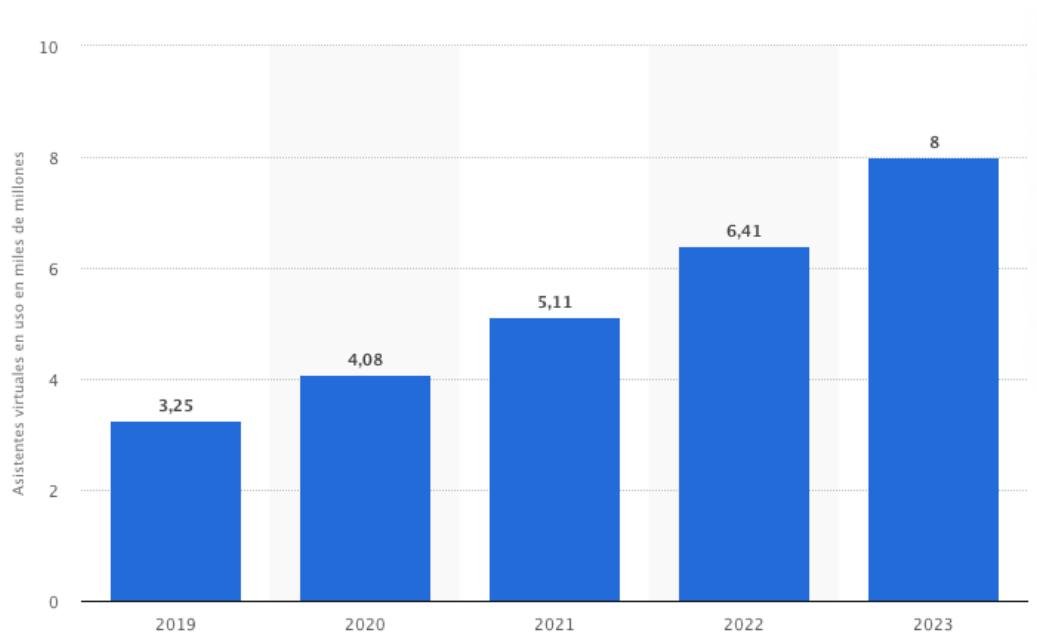


Ilustración 1.1. Miles de millones de AVI en uso por años

Otra de las motivaciones en el desarrollo de este proyecto es la realización del mismo en el idioma español, segunda lengua más hablada del mundo² y tercera lengua más utilizada en Internet³ como se muestra en la Ilustración 1.2..

² <https://es.babbel.com/es/magazine/los-10-idiomas-mas-hablados-del-mundo>

³ <https://www.internetworldstats.com/stats7.htm>

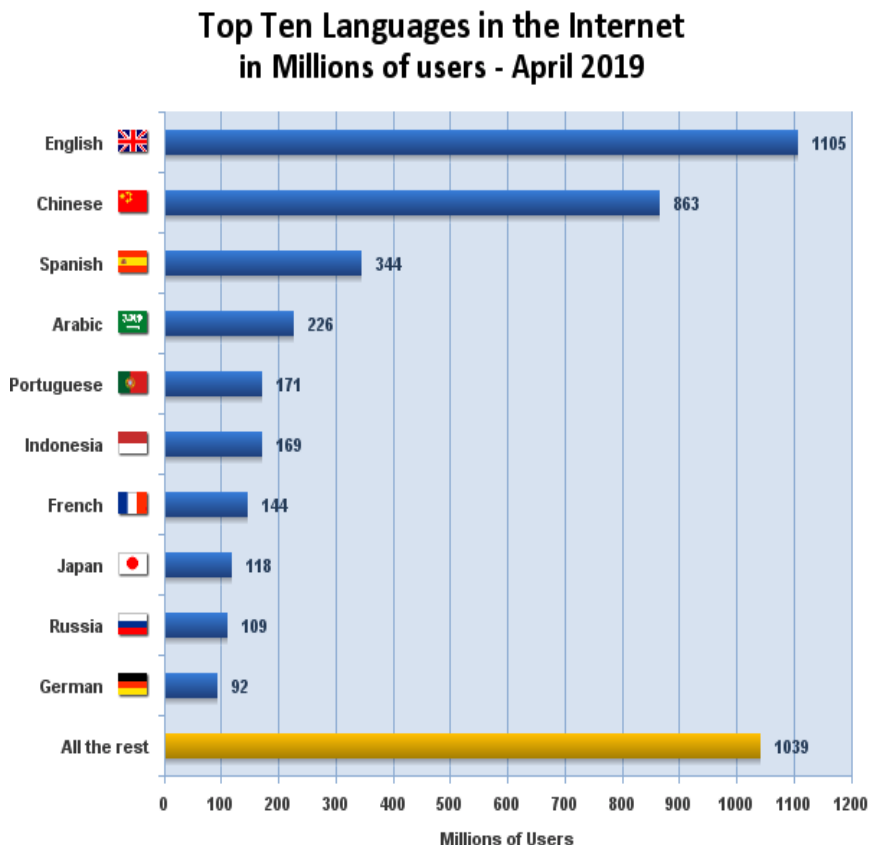


Ilustración 1.2. Millones de usuarios de Internet por lenguas.

Gran parte de los recursos existentes que ayudan a clasificar las emociones están en inglés, por lo que una de las tareas a realizar será la de estudiar los recursos que existan en español para determinar cuales pueden ayudar a realizar la tarea.

Este proyecto también me aportará un aprendizaje de nuevas técnicas que no se han desarrollado en profundidad durante el máster. Como primeras tareas realizaré una amplia investigación sobre desarrollo de utilidades para AVI y sobre la clasificación de las emociones, para posteriormente poder desarrollar la parte más técnica y práctica del proyecto mediante el diseño y la implementación de una utilidad, llamada skill, para Amazon Alexa capaz de reconocer la emoción que expresa el usuario y dar una respuesta acorde a la misma.

1.3. Propósito del proyecto

El propósito principal de este proyecto es llevar a cabo un estudio y su posterior desarrollo de una utilidad (skill) de un AVI, en concreto para Amazon Alexa, que incorporará un detector de emociones en función de lo que diga el usuario, para lo que

se aplicarán técnicas de PLN. Se analizarán distintos perfiles de usuario para establecer respuestas adecuadas en función de la emoción detectada.

1.4. Objetivos del proyecto

El principal objetivo del proyecto es crear una skill para el AVI de Amazon, Alexa, que sea capaz de reconocer la emoción que expresa un usuario al comunicarse con el agente. La skill será capaz de contestar de una forma u otra en función de la emoción que haya detectado. Esta meta puede dividirse en varias tareas, que se enumeran a continuación:

1. Realizar un estudio de las distintas soluciones de desarrollo de skills para AVI.
2. Estudiar la bibliografía y técnicas relacionadas con la clasificación de emociones.
3. Implementación del front-end de la Skill. Parte encargada de interactuar con el usuario, abrir la skill, pedir una frase para obtener la emoción y devolver la respuesta apropiada.
4. Implementar el detector de emociones incorporando alguna de las técnicas de clasificación de emociones estudiadas. Este formará el back end de la skill, es decir, la parte con la que el usuario no interactúa, se encarga de procesar los datos de entrada.
5. Realizar un conjunto de experimentos para comprobar los resultados que proporciona el sistema creado.
6. Redactar una memoria que recoja todo el trabajo desarrollado, así como los manuales de instalación y de usuario.

1.5. Descripción del problema

A continuación se explica la descripción del problema que se quiere solucionar mediante el desarrollo de este trabajo.

Actualmente el enorme desarrollo que está teniendo la IA, está provocando un desarrollo en los AVI, haciendo que surjan nuevos dispositivos inteligentes y que todos ellos estén experimentando un cambio hacia una mayor funcionalidad.

Este desarrollo que están teniendo los AVI está provocando que se creen multitud de nuevas habilidades, lo que está provocando que estos dispositivos se conviertan en un gran complemento para el día a día de las personas, llegando incluso a ser capaces de mantener una conversación lógica y normal con el usuario.

Por otro lado el PLN es una tarea puntera dentro del sector de la informática, ya que permite utilizar la enorme cantidad de datos que genera la población mundial a diario, en la que expresan sus emociones, su opinión acerca de un tema, etc. El PLN, como ya hemos dicho se ayuda de esta información y obtiene datos de ella, como por ejemplo cómo se encuentra de ánimo una persona, que piensa sobre una noticia, etc. Estos datos pueden ayudar en diversas ramas como la inteligencia de comercio, encuestas masivas, incluso en delitos informáticos.

Para utilizar estas dos grandes áreas de la informática, se va a desarrollar este proyecto, en el que se generará una Skill que sea capaz de reconocer la emoción que expresa el usuario al comunicarse con el AVI. Esta habilidad se desarrollará para Alexa de Amazon, que es uno de los más utilizados a nivel mundial.

1.6. Resultados esperados

Una vez finalizado el desarrollo del proyecto, los resultados que deben obtenerse son:

1. Investigación acerca de las diferentes opciones que existen para el desarrollo de skills.
2. Elección de técnica de reconocimiento de emociones.
3. Sistema capaz de reconocer emociones a partir de un texto de entrada.
4. Skill para Amazon Alexa capaz de dar una respuesta acorde al estado de ánimo.
5. Memoria del proyecto
6. Manuales de instalación y de usuario.

1.7. Planificación

El objetivo de la planificación de un proyecto es ajustar un ámbito de trabajo, de manera que facilite estimar de manera razonable el tiempo y los costes que serán

necesarios para el desarrollo del proyecto. Estas estimaciones son flexibles y pueden ir adaptándose a la situación del proyecto.

1.7.1. Estimación temporal

La planificación temporal se llevará a cabo mediante un diagrama de Gantt, que se trata de una herramienta gráfica que suele ser utilizada para mostrar el tiempo que se le va a dedicar a un conjunto de tareas dentro de un marco temporal determinado.

En la siguiente tabla (Tabla 1.1) se muestran las tareas en que se ha dividido el proyecto junto con la duración y las fechas asociadas a cada una de ellas:

Tarea	Duración	Fecha inicio	Fecha fin
Estudio, comparación de los diferentes AVI. Elección de uno de ellos para el desarrollo de la skill.	18 días	18/11/2019	05/12/2019
Estudio de las diferentes técnicas para la clasificación de emociones	21 días	03/12/2019	23/12/2019
Implementación de un sistema de reconocimiento de emociones utilizando la técnica de clasificación elegida	36 días	24/12/2019	28/01/2020
Diseño e implementación de la skill que incorpora el sistema de reconocimiento de emociones	43 días	29/01/2020	11/03/2020
Realización de pruebas para la experiencia de usuario, así como para asegurar la estabilidad y la seguridad del sistema	9 días	12/03/2020	20/03/2020
Desarrollo de los manuales de usuario y de instalación	10 días	23/03/2020	01/04/2020
Generación de una memoria que contenga todo el trabajo realizado	167 días	01/12/2019	15/05/2020

Tabla 1.1. Estimación temporal de las tareas del proyecto.

Estas tareas llevan asignado un tiempo limitado por una fecha inicial y una final, aunque estas fechas pueden variar a medida que se desarrolle el proyecto.

A continuación, se muestra el gráfico mencionado anteriormente (Ilustración 1.3.). En él se pueden observar las tareas listadas junto con la duración de cada una de ellas. Este gráfico se ha creado mediante una web online⁴. El periodo de tiempo que se estima que ocupe la realización de este proyecto es de 6 meses.

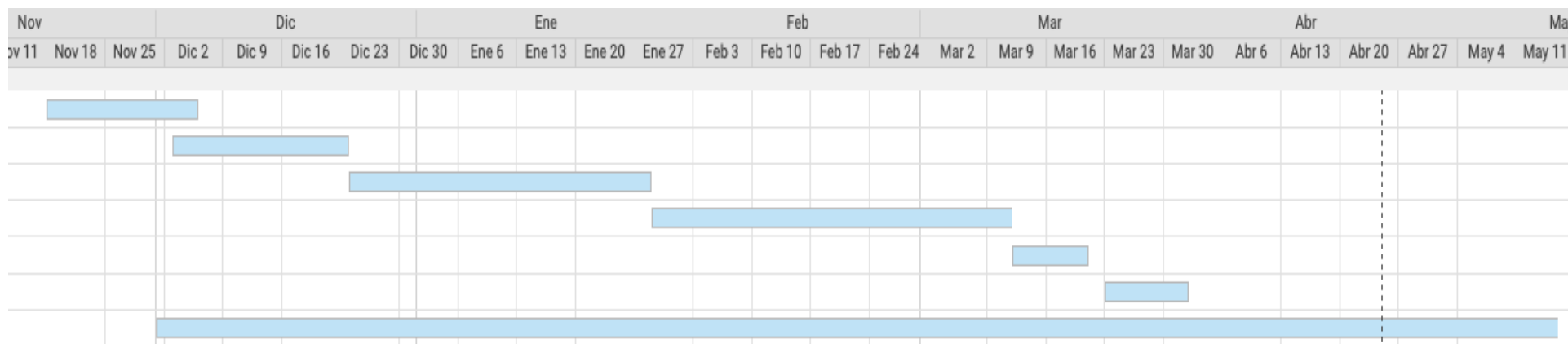


Ilustración 1.3. Diagrama de Gantt

⁴ <https://es.smartsheet.com/software-diagrama-gantt>

1.7.2. Estimación de costes

Todo proyecto software genera una serie de costes durante su desarrollo y su posterior mantenimiento. Para el cálculo de estos costes hay que tener en cuenta diferentes costes, como son los costes hardware, costes software, costes de personal y otros costes. Para el cálculo de estos costes se tendrá en cuenta el tiempo que durará el desarrollo del proyecto.

1.7.2.1. Costes hardware

Los costes hardware representan los costes asociados al equipo físico necesario para el desarrollo del proyecto. A continuación, se muestran las especificaciones utilizadas para el desarrollo del proyecto.

- Ordenador portátil: Macbook Pro 1580€
 - Procesador: 2,3 GHz Intel Core i5 de doble núcleo.
 - Memoria RAM / SDD: 8GB / 256GB.
 - Tarjeta gráfica: Intel Iris Plus Graphics 640 1536 MB.
 - Pantalla de retina de 13.3" (2560 x 1600 pixeles).

El coste mostrado se debe amortizar durante el desarrollo del proyecto. Se estima una vida útil de 8 años, pasado este tiempo, el valor será de 0€. Obtenemos un coste de 197,5€ al año, lo que equivale a 16,5€ al mes. Dado que el proyecto tiene una duración estimada de 6 meses el coste de hardware será de 99€.

1.7.2.2. Costes software

Los costes software engloban los gastos provocados por la adquisición de diversas licencias de programas necesarios para la realización del proyecto.

Los programas utilizados y que producen los costes de software son los siguientes:

- Sistema operativo macOS 0€
- Navegador web 0€
- Google Drive 0€
- Microsoft Office 2016 69€/año

Estos costes software también se tienen que amortizar durante el desarrollo del proyecto. Teniendo en cuenta que el software de Microsoft Office 2016 se utilizará durante 5 meses y medio, su coste sería de 31,62€, que será el coste total de software.

1.7.2.3. Costes de personal

Los costes de personal representan los producidos por los sueldos que habría que pagar para la contratación de personal para el desarrollo del proyecto. En este caso será realizado por una única persona, aunque tendrá diferentes roles dentro del desarrollo del mismo. Cada uno de ellos tendrá unos costes diferentes.

Se necesitan los siguientes perfiles de personal:

- **Analista:** se encargará de analizar el sistema, diseñarlo e intentar hacerlo más eficiente posteriormente.
- **Programador:** se encargará de implementar el sistema.
- **Técnico:** se encargará de realizar los diferentes manuales para el uso del sistema, realizar pruebas y evaluar el sistema.

Los salarios anuales de cada persona se han obtenidos del BOE del 19 de Junio de 2013⁵. El máximo de horas de trabajo semanales también se ha extraído de la misma web y es de 39 horas/semana.

Para la realización del proyecto se estima un total de 360 horas, lo que nos lleva a un coste de personal que será mostrado en la siguiente tabla (Tabla 1.2.).

Personal	Salario anual	Salario/hora	Nº horas	Total
Analista	25.830€	14,16€	90	1274,4€
Programador	20.734€	11,37€	240	2728,8€
Técnico	15.554 €	8,53€	30	255,9€

Tabla 1.2. Costes de personal

⁵ https://www.boe.es/diario_boe/txt.php?id=BOE-A-2013-6678

El coste total de personal será de 4259,1€.

1.7.2.4. Otros costes

Otros costes que no han sido tenidos en cuenta hasta el momento son los producidos por la conexión a Internet, que tendrá un coste de 12,84€/mes y un total de 100MB con Vodafone.

El coste total será el de la conexión durante los 6 meses que durará el proyecto, es decir, 77,04€

1.7.2.5. Coste total

El coste total del proyecto estará formado por la suma de todos los costes explicados anteriormente (costes de hardware, costes de software, costes de personal y otros costes). En la siguiente tabla (Tabla 1.3.) se puede observar un resumen de todos los gastos generados por el proyecto.

Concepto	Coste
Costes hardware	99€
Costes software	31,62€
Costes de personal	4259,1€
Otros costes	77,04€
TOTAL	4466,76€

Tabla 1.3. Resumen de gastos totales del proyecto.

Además de estos costes se considera aplicar un 10% de beneficio, por lo que el coste total será de **4913,436€**.

1.8. Estructura del proyecto

A continuación, se explica la estructura que forma la memoria de este proyecto, la cual está formada por cinco capítulos:

- En el capítulo 1, titulado “Introducción”, se explican los motivos y propósitos del proyecto junto con los objetivos esperados y los resultados obtenidos. También se presentan los costes y la planificación temporal.
- En el capítulo 2, titulado “Asistentes virtuales inteligentes”, se lleva a cabo una explicación acerca de los asistentes virtuales, los más importantes, sus usos y una comparativa de ellos.
- En el capítulo 3, titulado “Reconocimiento de emociones”, se explican los diferentes modelos de emociones, técnicas para el reconocimiento de emociones y los recursos utilizados.
- En el capítulo 4, titulado “Desarrollo del proyecto”, como su propio nombre indica se explica el proceso para la creación del proyecto. Primero, se detalla la descripción, los objetivos y requerimientos del proyecto. Posteriormente, se elige una metodología de Ingeniería del Software y se enumeran las diferentes historias de usuario que forman el proyecto. Por último, se detalla el desarrollo de cada una de las historias de usuario y se comentan las diferentes tecnologías utilizadas para el desarrollo del sistema.
- En el capítulo 5 y último, titulado “Conclusiones y trabajos futuros”, se explican las conclusiones y valoraciones obtenidas tras la realización del proyecto, además de los posibles trabajos futuros a desarrollar para mejorar el sistema.
- En el apartado titulado “Bibliografía”, se indican los recursos que se han consultado para la realización de este proyecto y referencias a los mismos.
- Por último, esta memoria también incluye cuatro anexos:
 - Anexo A , titulado “Manual de instalación”, contiene las instrucciones necesarias para poder utilizar la herramienta creada durante el desarrollo de este trabajo.
 - Anexo B, titulado “Manual de usuario”, contiene un breve manual que ayude al usuario a utilizar la skill que se ha creado.

- Anexo C, titulado “Índice de ilustraciones”, contiene el índice de las ilustraciones que aparecen en el proyecto.
- Anexo D, titulado “Índice de tablas”, contiene el índice de las tablas que aparecen en el proyecto.

Capítulo 2. Asistentes virtuales inteligentes

En este capítulo se va a desarrollar que es un AVI, junto con los sistemas conversacionales, explicando sus diferencias y similitudes. Posteriormente se presentarán los AVI más potentes del mercado, centrándonos en Alexa, para el que se va a llevar a cabo este proyecto. Por último, se realizará una comparativa de los AVI explicados previamente.

2.1 Asistentes virtuales inteligentes y sistemas conversacionales

Un AVI⁶⁷ es un agente basado en software que tiene habilidades de aprendizaje permanente y cuyo objetivo consiste en ayudar a los usuarios de sistemas computacionales a la hora de realizar multitud de tareas, haciendo que la interacción entre la persona y la máquina sea mínima.

Por otro lado, la interacción entre el asistente (normalmente representado por una voz femenina) y el usuario tiene dos pasos principales:

1. El usuario se comunica utilizando la voz con el asistente.
2. Posteriormente el agente procesa el mensaje detectado del usuario, lo interpreta y procede a dar una respuesta coherente, ayudándose en la mayoría de los casos de una gran variedad de recursos online (clima, noticias, etc.) y de su propia inteligencia, que va aumentando a medida que interactúa con el usuario.

Algunas de las tareas que son capaces de realizar son:

- Llamar a un contacto.
- Enviar un mensaje.
- Poner una alarma.
- Realizar una búsqueda en la web.
- Poner canciones en plataformas como Spotify.
- Crear eventos en el calendario.

⁶ https://es.wikipedia.org/wiki/Asistente_virtual

⁷ <http://tublogtecnologico.com/los-asistentes-virtuales-dispositivos-pueden-hacerte-la-vida-mas-comoda/>

Por otro lado, un sistema conversacional⁸⁹ es un programa que permite al usuario simular una conversación con él mediante el lenguaje natural (voz o texto). Estos sistemas se caracterizan por ser menos inteligentes que los asistentes virtuales ya que no siempre entienden el mensaje enviado por el usuario. Se basan en una serie de respuestas preparadas previamente. Estos sistemas no saben de qué están hablando realmente, por lo que la conversación que se tiene con el usuario puede tener más o menos lógica.

2.2. Principales asistentes virtuales inteligentes

En este apartado se van a explicar los AVI más importantes en el mercado¹⁰¹¹ actualmente, explicando su nacimiento, sus principales características y modos de uso.

2.2.1. Siri

Siri¹² es el AVI de Apple, creado por Dag Kittlaus, Adam Cheyer y Tom Gruber en Diciembre de 2007. A España llegó junto con iOS 6 en Septiembre de 2012. Una de sus características destacables son sus respuestas graciosas y hasta filosóficas.

Nació como una aplicación independiente, hasta que posteriormente fue integrado en los dispositivos de Apple. Siri es capaz de aprender sobre nuestros gustos personales para realizar un servicio más personalizado. Únicamente permite comunicarnos mediante la voz con ella.

2.2.2. Google Assistant

Google Assistant¹³ se trata del asistente virtual desarrollado por Google, que nos ayuda a realizar diversas tareas como obtener respuestas a nuestras preguntas, planificar nuestras tareas diarias y controlar nuestro hogar mediante el asistente Google Home. Fue dado a conocer en Mayo de 2016. A diferencia de Google Now, es capaz de mantener conversaciones bidireccionales. Se activa al escuchar “OK google”

⁸ https://es.wikipedia.org/wiki/Bot_conversacional

⁹ <https://planetachatbot.com/diseño-de-un-sistema-conversacional-nivel-prototipo-con-google-actions-a44ceb4a5a6f>

¹⁰ <https://www.cnet.com/es/noticias/siri-alex-google-assistant-asistentes-virtuales-cual-nos-gusto-mas/>

¹¹ <https://contactcenterhub.es/los-asistentes-virtuales-revolucionan-el-mercado-2019-19-19013/>

¹² <https://www.apple.com/es/siri/>

¹³ <https://assistant.google.com>

y está disponible para móviles Android (donde va integrado) y iPhone (donde hay que descargar la aplicación desde App Store), pantallas inteligentes, televisores con Android TV, smartwatch wear OS y altavoces inteligentes. Google Assistant nos permite comunicarnos con él mediante texto o mediante voz natural.

2.2.3. Cortana

Cortana¹⁴ es el AVI desarrollado por Microsoft. Nació el 2 de Abril de 2013. Está integrado en Español en Windows 10 y Windows Phone. Es similar a Siri, con la ventaja de que desde el primer momento se le puede introducir una serie de datos personales con los que podrá darnos consejos o sugerencias similares a nuestros gustos. Además, irá conociéndonos más a fondo con el tiempo y mejorando sus respuestas. Cortana puede ser utilizada mediante voz y también mediante texto.

2.2.4. Bixby

Bixby¹⁵ es el AVI creado por Samsung, nació el 20 de Marzo de 2017, cuando solo estaba disponible en el Samsung Galaxy S8 y S8+. Actualmente su principal misión consiste en conseguir que hablemos con él con total naturalidad y que todo lo que podamos hacer mediante una iteración, podamos hacerlo mediante la voz. Permite llevar a cabo las tareas básicas de un asistente virtual (gestionar agenda, configurar alarmas, etc.). Bixby analiza nuestros hábitos para ayudarnos en nuestra rutina diaria. Permite crear comandos rápidos mediante la asignación de una palabra o frase clave y asociarla a una acción.

2.2.5. Alexa

Alexa¹⁶, es el AVI propiedad de Amazon. Nació en Noviembre de 2014. Es posible utilizarlo no solo en sus dispositivos Echo, sino también en cualquier dispositivo mediante su aplicación ("Amazon Alexa"). Alexa fue el primer asistente que se vio reflejado en un producto físico (Amazon Echo). El motor back-end en el que se basa se ejecuta en la nube de Amazon Web Services.

¹⁴ <https://www.microsoft.com/es-es/windows/cortana>

¹⁵ <https://www.samsung.com/es/apps/bixby/>

¹⁶ <https://developer.amazon.com/es-ES/alexa>

El asistente se activa simplemente con escuchar su nombre, y nos ayuda a realizar diversas tareas y solucionar dudas. En cuanto a su integración en un hogar, es capaz de realizar tareas como encender las luces de un dormitorio o poner una lavadora.

Alexa recolecta diariamente multitud de datos de sus usuarios ya que procesa una enorme cantidad de interacciones cada día, lo que le permite ir aprendiendo acerca de los gustos y rutinas de sus usuarios, haciendo así más simples las respuestas a las interacciones.

Este ayudante está conectado con diversos servicios que nos ayudarán a hacer multitud de actividades como por ejemplo pedir una pizza de Domino's. También es posible desarrollar "funcionalidades", llamadas por Amazon "Skills", que los usuarios pueden crear de forma libre y gratuita.

A la hora de utilizar Alexa, existen dos posibilidades, que se muestran en la Ilustración 2.1.

- Alexa Voice Server, que permite a las compañías incorporar Alexa en sus propios dispositivos.
- Alexa Skills Kit (ASK), que permite a los desarrolladores extender la funcionalidad de Alexa.



Ilustración 2.1. Formas de utilizar Alexa

En este TFM vamos a usar esta utilidad para realizar nuestra Skill. ASK es un conjunto de herramientas, ejemplos de código, documentación y API que permite crear Skill para Alexa.

Las dos partes que forman una skill (Ilustración 2.2.) y que hay que desarrollar para crearla son:

- **Interfaz de usuario de voz (front-end):** se trata de lo que el skill debe de comprender cuando el usuario le hable.
- **Lógica de programación (back-end):** consiste en el tratamiento que debemos de hacer para dar una respuesta para que Alexa conteste al usuario. Es posible realizarlo mediante ASW Lambda (en Python o Node.js) o llevarlo a cabo por otro lado independiente (en este caso se debe de unir con el front-end).



Ilustración 2.2. Partes de una skill

A la hora de ejecutar una Skill de Alexa, la orden debe tener una serie de componentes, que se muestran en la Ilustración 2.3. y se explican a continuación.



Ilustración 2.3. Orden para ejecutar una skill de Alexa

- Palabra para despertar (wake word): es la palabra que hace que el asistente se active y comience a escuchar. Por defecto es “Alexa”, aunque se puede cambiar por otras como “echo” o “Amazon”.
- Palabra de lanzamiento (launch): se trata de la palabra que indica que debe de hacer. Ej.: abre, lanza, comienza, pregunta, etc.
- Nombre de la skill que queremos abrir para utilizar.
- Enunciado (utterance): indican la acción concreta dentro del Skill que queremos que haga. Ej.: dame un dato, cuéntame algo, dime hola, etc. Es posible añadir una “variable” al enunciado mediante el uso de “Slots”, que son variables (mes, día, etc) con un tipo asignado. Para crear un “Intent” (habilidad que es capaz de hacer Alexa) debemos proporcionar varios enunciados que el usuario pueda

decir para referirse al mismo. No es necesario pensar todas las posibilidades ya que Alexa puede aprender. Ej.: Para un intent que sea decir hola mundo ("helloWorldIntent"), podemos crear varios enunciados como "di hola" "hola mundo" "di hola mundo" "cómo estás", etc. Al escuchar cualquiera de esos enunciados Alexa sabe que queremos ejecutar el Intent "helloWorldIntent".

Existen una serie de Intent predeterminados (cancelar, ayuda y parar) que deben de ser añadidos a nuestra skill para que sea verificada. En la siguiente ilustración (Ilustración 2.4) se muestra la estructura de intent y utterances.

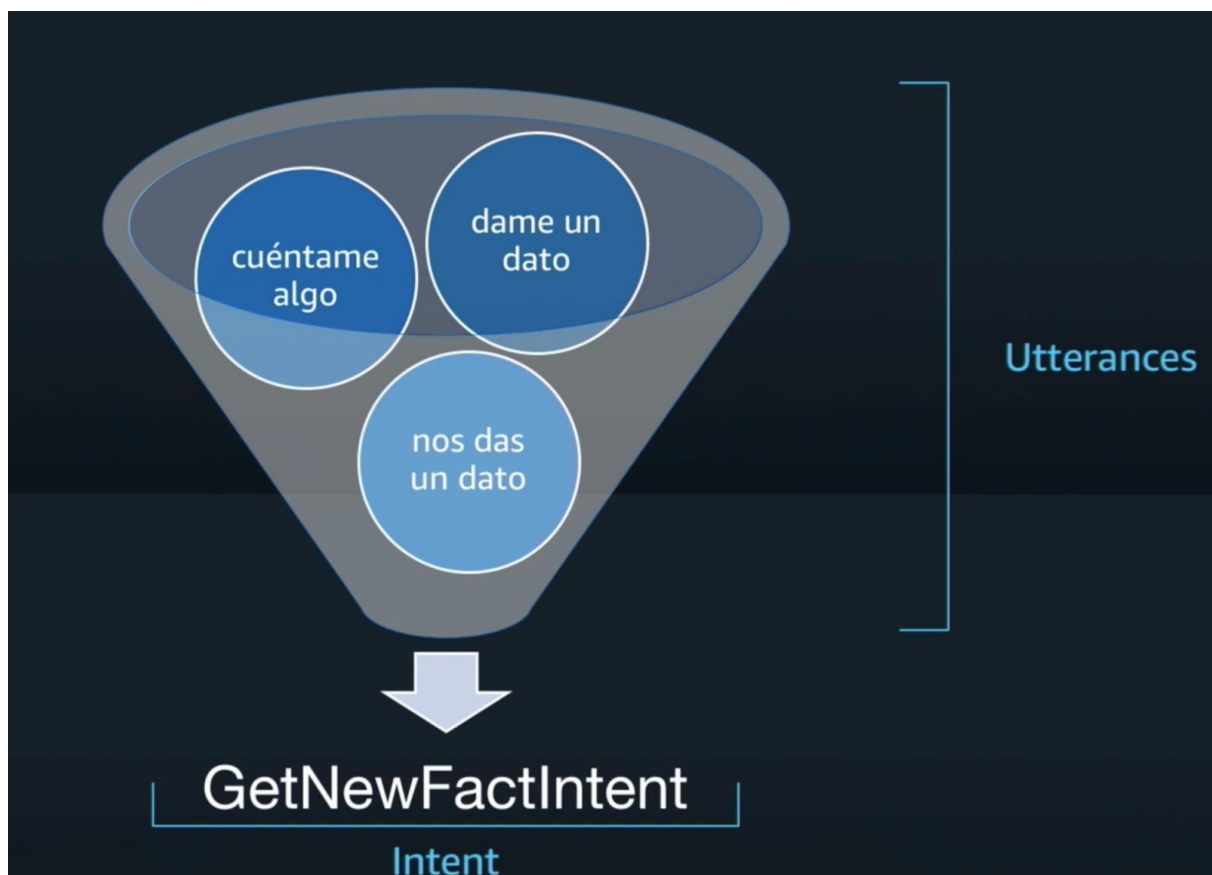


Ilustración 2.4. Intent y utterances

El funcionamiento de Alexa cuando un usuario le envía una orden es el siguiente (Ilustración 2.5.):

1. El usuario dice la palabra para despertar el dispositivo, éste se activa, comienza a escuchar y captura el mensaje que el usuario envía.
2. Este mensaje se envía al servicio de Alexa donde se le aplican dos pasos principales:

- a. Automatic Speech Recognition (ASR): reconocimiento automático del audio y transformación a texto del mismo.
 - b. Natural Language Understanding (NLU): se lleva a cabo un proceso de machine learning y de procesamiento del lenguaje natural para entender lo que el usuario nos quiere decir y lo que quiere que haga el asistente.
3. Una vez tenemos el Intent que el usuario quiere ejecutar se envía al back-end de la skill. En él deberá decidirse lo que hacer para dar una respuesta al servicio de Alexa.
 4. Por último, el texto enviado por nuestro back-end se recibe en el servicio de Alexa donde se lleva a cabo una última etapa de machine learning que consiste en transformar el texto a habla.

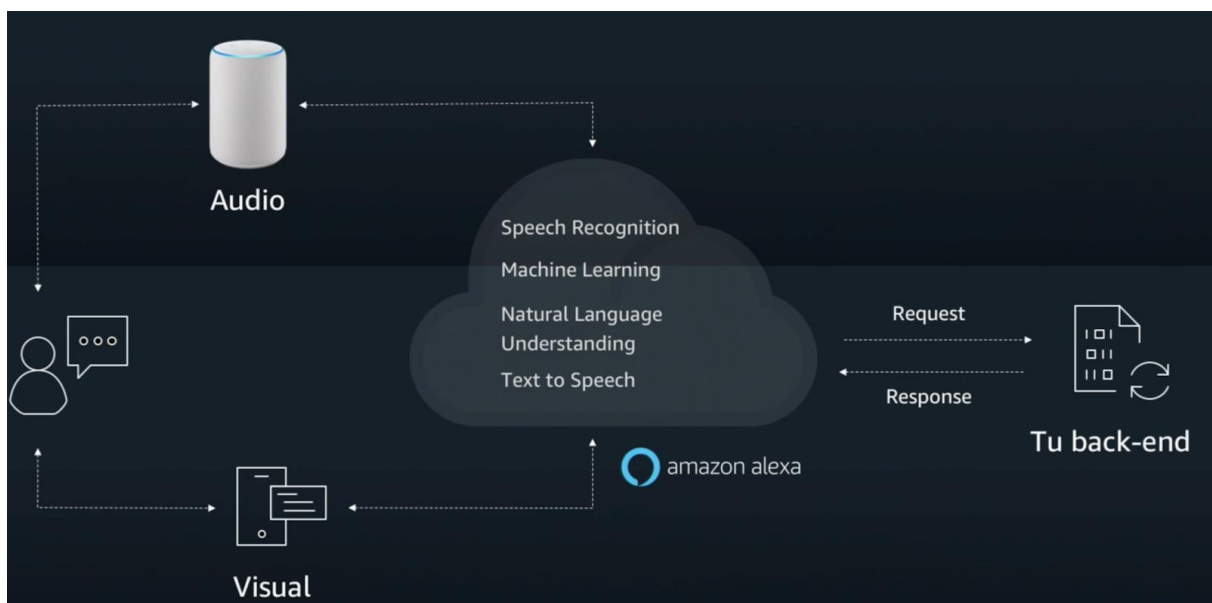


Ilustración 2.5. Funcionamiento de Alexa

Para el futuro, Amazon pretende cambiar el enfoque, que actualmente consiste en interacciones pasivas, ya que Alexa no se activa mientras no se le reclame, hacia un enfoque de proactividad, en el que Alexa se anticipe a nuestras necesidades¹⁷, ayudándose para ello de todos los datos recolectados sobre el usuario.

¹⁷ <https://www.xataka.com/inteligencia-artificial/proximo-paso-evolucion-alexa-proactividad-asistente-amazon-usara-ia-para-anticiparse-a-nuestros-deseos>

2.3. Comparativa asistentes virtuales inteligentes

La firma de capitales “Loup Ventures”¹⁸, especializada en la realización de diversos análisis sobre IA y realidad virtual llevó a cabo un estudio¹⁹ en el que se comparan la capacidad de comprensión y el rendimiento de sus respuestas de los cuatro asistentes virtuales más importantes del mercado:

- Google Assistant
- Siri
- Alexa
- Cortana

En el estudio se hicieron un total de 800 preguntas de diversos temas a cada asistente, obteniendo los siguientes resultados (Tabla 2.1.).

	Respondió correctamente	Consulta entendida
Google Assistant	87.9%	100%
Siri	74.6%	99.6%
Alexa	72.5%	99.0%
Cortana	63.4%	99.4%

Tabla 2.1. Comparativa funcionamiento AVIs

Mediante estos resultados se confirma a Google Assistant como el asistente virtual más inteligente, ya que comprende el 100% de las preguntas que se le proponen (siendo el más eficaz en cuatro de las cinco categorías en que se dividieron las preguntas) y responde correctamente más de un 10% de preguntas respecto a su perseguidor, Siri.

Los resultados obtenidos en la 1ª Ola del Estudio General de Medios de 2019²⁰ (EGM), para la Asociación para la Investigación de Medios de Comunicación (AIMC) muestran (Ilustración 2.6.) que el asistente virtual más utilizado en España es Siri (Apple), seguido por el Asistente de Google y Alexa (Amazon) en tercera posición.

¹⁸ <https://loupventures.com>

¹⁹ <https://loupventures.com/annual-smart-speaker-iq-test/>

²⁰ https://www.aimc.es/a1mc-c0nt3nt/uploads/2019/04/190404_NP_EGM_2019ola1.pdf

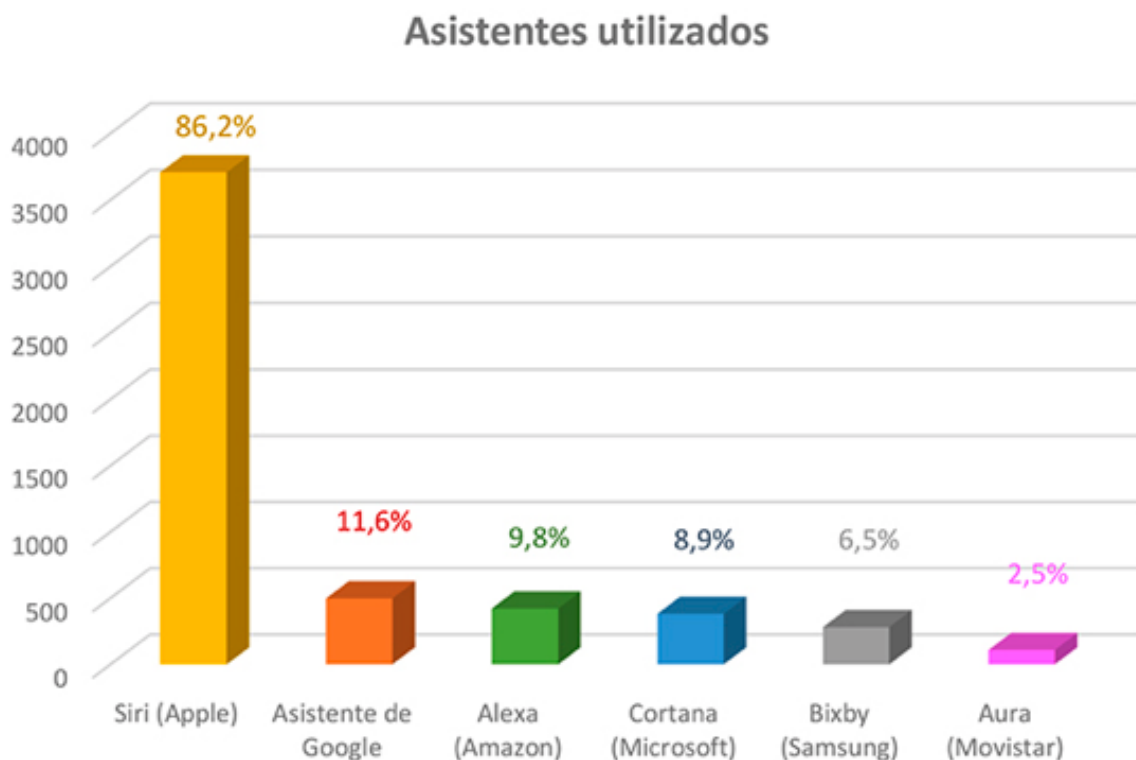


Ilustración 2.6. Asistentes virtuales más utilizados en España

En ese mismo estudio se afirma que un 10,7% de la población de España (4.300.000 individuos) hace uso de los asistentes virtuales de voz a diario.

Por otro lado, los altavoces inteligentes que contienen un asistente virtual de los vistos anteriormente más presentes en los hogares españoles son Google Home y Amazon Alexa, como se muestra en la Ilustración 2.7.

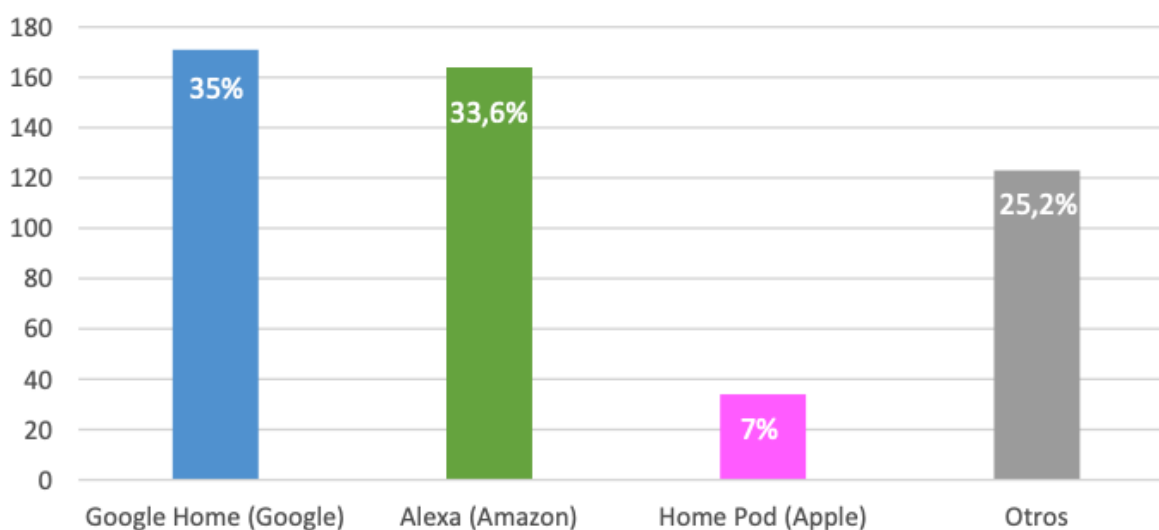


Ilustración 2.7. Dispositivos con AVI más utilizados

En cuanto al dispositivo desde el que los usuarios utilizan estos asistentes virtuales, el portátil y el móvil son los más utilizados, acaparando un 87.2%, como se muestra en la siguiente ilustración (Ilustración 2.8).

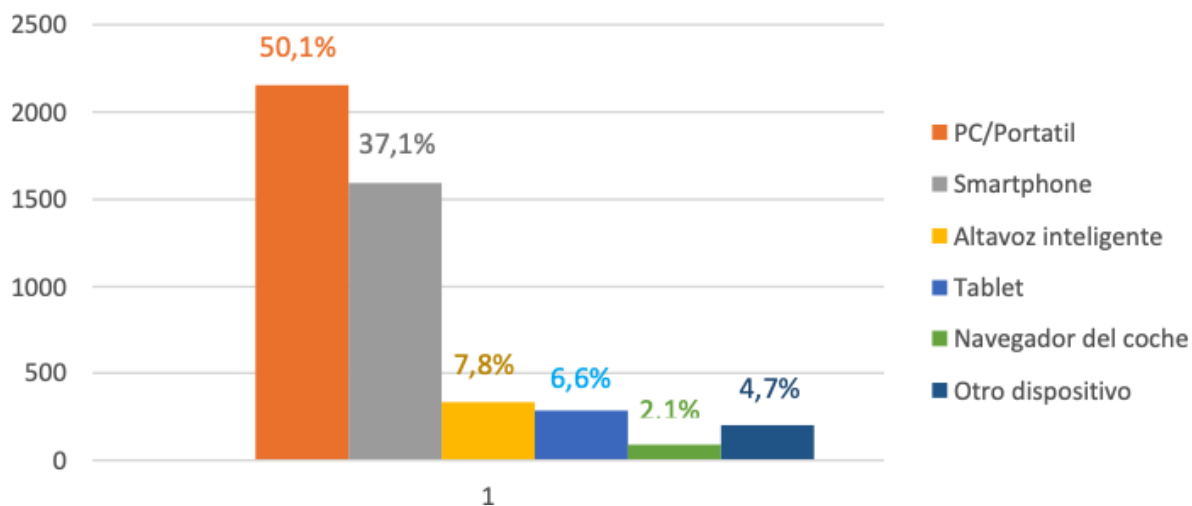


Ilustración 2.8. Dispositivos desde donde usuarios utilizan un AVI

Por tanto, se ha decidido realizar el proyecto para el AVI Alexa de Amazon, ya que es uno de los más utilizados en España (donde se habla español que es el idioma en que se desarrollará el proyecto). Además Amazon proporciona multitud de material y recursos para ayudar a sus desarrolladores en la creación de nuevas skills.

Capítulo 3. Reconocimiento de emociones

En este capítulo se desarrollan varios aspectos del reconocimiento de emociones, los diferentes modelos de emociones que existen, las diferentes técnicas que se pueden utilizar para dicha tarea, y por último se explicarán los lexicones que se han estudiado y que finalmente se han decidido utilizar para la realización de este proyecto.

Antes de proceder a desarrollar los puntos citados anteriormente, es conveniente explicar el marco en el que se encuentra el reconocimiento de emociones.

Las emociones son rasgos básicos de los humanos y ya han sido estudiadas por investigadores de diversos campos como la psicología, la medicina, la sociología y como no por la informática.

El reconocimiento de emociones se enmarca dentro de un área llamada Computación Afectiva, la cual es muy amplia y abarca diversas disciplinas. Una de ellas, en la que nos centraremos, es el Análisis de Sentimientos [10], que abarca todas las tareas encargadas de la detección, el análisis y la evaluación del estado anímico de los humanos ante diferentes situaciones.

Las diversas tareas que se enmarcan dentro del Análisis de Sentimientos se muestran en la siguiente ilustración (Ilustración 3.1.) y se explicarán a continuación.

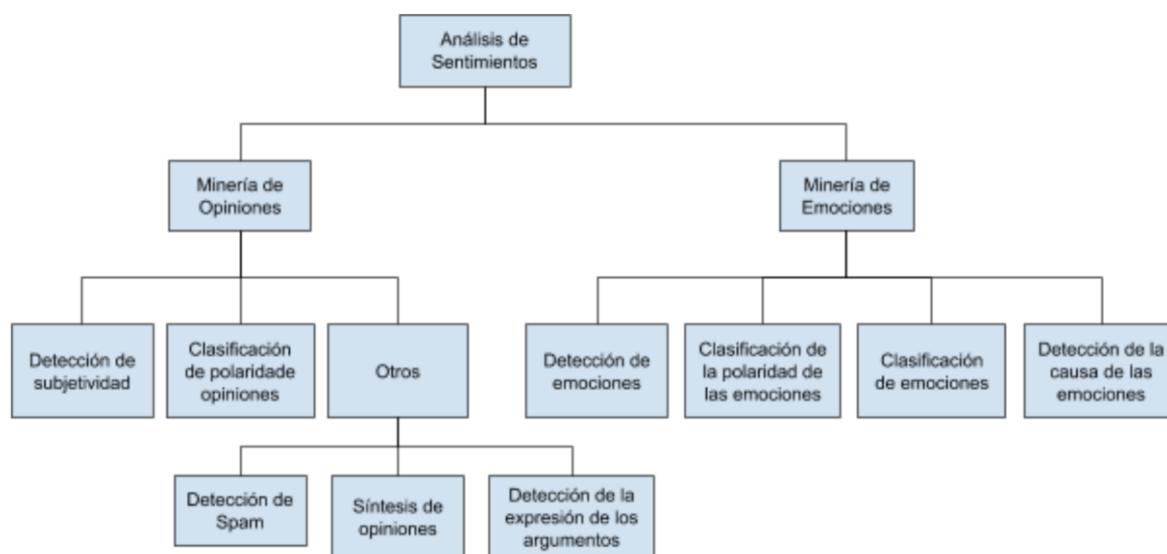


Ilustración 3.1. Tareas dentro del Análisis de sentimientos

Existen dos grandes tareas dentro del Análisis de Sentimientos: minería de opiniones y minería de emociones. A continuación, se explican las subtareas de la minería de emociones, ya que es la tarea que contiene la clasificación de emociones que es el propósito a desarrollar en este proyecto.

- Detección de emociones: trata de detectar si un texto transmite alguna emoción determinada o no.
- Clasificación de la polaridad de las emociones: trata de extraer de un texto, que se asume que contiene alguna emoción, la polaridad de la misma.
- Clasificación de emociones: consiste en clasificar las emociones existentes en un texto dentro de un conjunto de emociones predefinido.
- Detección de la causa de las emociones: trata de extraer los factores que pueden provocar los diferentes tipos de emociones.

En este trabajo nos centramos en la tarea de la clasificación de emociones, con el objetivo de reconocer las diversas emociones que puede el usuario expresar diciendo una oración mediante la voz, y ser capaces de obtener la emoción que presenta más intensidad en dicha oración, la cual deberemos de transformar a texto para poder analizar. Una vez obtenida la emoción que se expresa, el asistente dará una respuesta al usuario acorde a su estado de ánimo.

3.1. Modelos de emociones

En el campo de la psicología, existen diversas teorías de clasificación de emociones, las cuales se dividen en dos vertientes principales: modelos de emociones discretas y modelos de emociones dimensionales [8][9]. A continuación, se explica cada una de las vertientes y algunos ejemplos de cada una de ellas.

3.1.1. Modelos de emociones discretas

Esta teoría defiende que existe un número limitado de emociones que una persona puede expresar o experimentar. Cada emoción surge de un sistema neuronal distinto. Se dividen las emociones en varios grupos y se asigna cada emoción a la categoría discreta que más se le ajuste. Los modelos pertenecientes a dicha teoría son los más utilizados debido a sus buenos resultados a pesar de su sencillez, ya que el espacio de las emociones está dividido en un número determinado de categorías

[9]. A continuación, se van a explicar dos de los principales modelos de emociones discretas:

- Ekman [11] propuso un modelo de clasificación de emociones basada en seis categorías (ira, asco, miedo, alegría, tristeza y sorpresa), las cuales se muestran en la Ilustración 3.2., relacionadas cada una de ellas con alguna expresión fisiológica facial, como se muestra en la siguiente ilustración. Este modelo es el más popular en el ámbito del PLN y en él se basan la mayor parte de los modelos pertenecientes a la vertiente discreta [9].



Ilustración 3.2. Emociones de Ekman²¹

Posteriormente creó una versión extendida [9] de su modelo de emociones donde añadió a las primeras seis las siguientes emociones: diversión, desprecio, satisfacción, orgullo por los logros, alivio, satisfacción, placer sensorial y vergüenza.

- Plutchik [12] definió la clasificación de las emociones en un total de ocho categorías. Las emociones que forman esta clasificación son las seis emociones básicas de Ekman y dos nuevas: confianza y anticipación. Estas ocho emociones se organizan en parejas bipolares, como se observa en la siguiente ilustración (Ilustración 3.3.): alegría y tristeza, ira y miedo, confianza y disgusto, sorpresa y anticipación. El uso de una emoción anula su opuesta. A partir de estas categorías es posible recrear un conjunto de

²¹ <https://emotionresearchlab.com/es/blog/expresiones-faciales-universales-y-emociones/>

emociones más complejo, denominado como “Rueda de las emociones de Plutchik”.

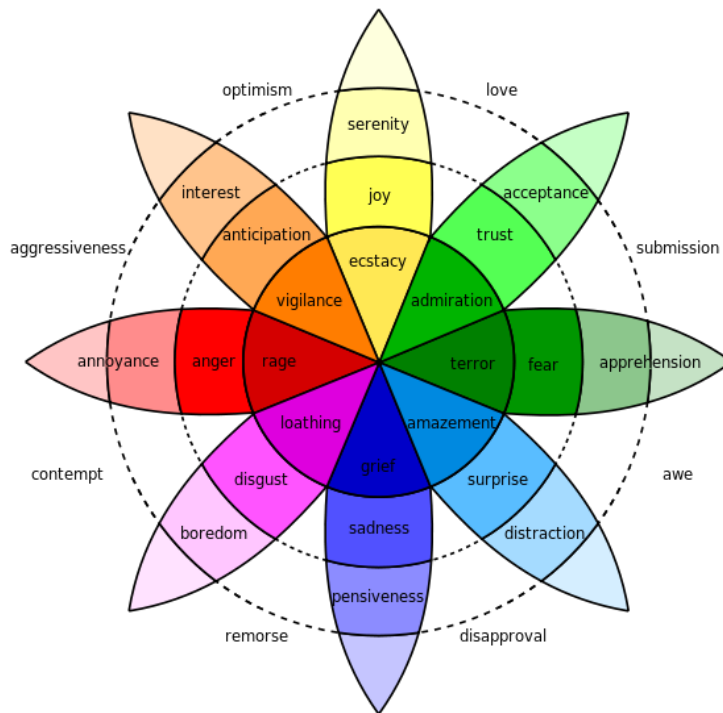


Ilustración 3.3. Emociones de Plutchik²²

3.1.2. Modelos de emociones dimensionales

Los modelos dimensionales [8][9] captan las emociones como un espacio con varias dimensiones donde la emoción que se exprese se representa como un punto en un espacio definido por diferentes variables que se tienen en cuenta, por lo que una emoción no es asignada a una única categoría. Esta teoría defiende que todas las emociones se encuentran en un espacio neurofisiológico donde en función de unos valores, cada emoción se encontrará en un punto diferente. La principal ventaja frente a los modelos discretos es que, al no tener un número de categorías definidas, cubre todas las emociones posibles. A continuación, se explican dos de los principales modelos pertenecientes a esta variante.

- Russell [13] defendía una clasificación de las emociones dentro de un espacio circular bidimensional, como se muestra en la Ilustración 3.4., con una dimensión para la valencia y otra para la excitación. La primera

²² <https://psicopico.com/la-rueda-las-emociones-robert-plutchik/>

dimensión indica el nivel de agrado o desagrado de la emoción mientras que la segunda indica el grado de activación de la emoción.

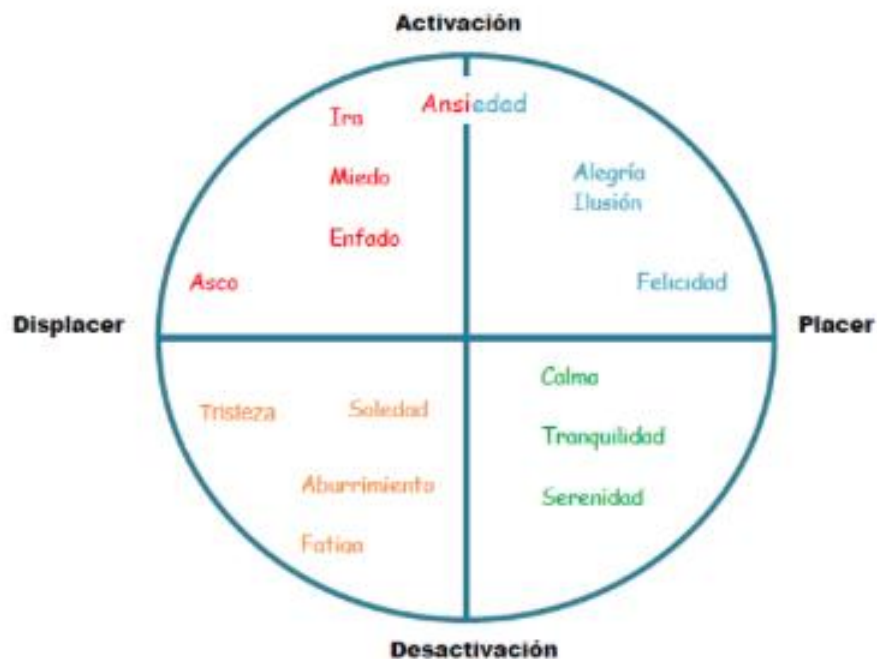


Ilustración 3.4. Emociones de Russell²³

- Mehrabian [14] creó un modelo basado en una representación en tres dimensiones independientes para describir el estado de ánimo de las personas. Este modelo es conocido como PAD, iniciales en Inglés de cada una de las dimensiones por las que se forma el espacio, que son:
 - Placer: concebido como un valor que podía ir desde el dolor extremo a la felicidad extrema.
 - Excitación: concebida como una actividad mental en una única dimensión que iba del sueño a la activación frenética.
 - Dominio: relacionada con los sentimientos de control y la medida en que una persona se ve restringido en su comportamiento.

Como ya se ha visto, existen multitud de opciones válidas para la clasificación de las emociones, pero no existe un consenso entre todas ellas. Por tanto, para la realización de este trabajo se ha decidido utilizar un modelo discreto para la clasificación de las emociones, ya que es uno de los modelos de emociones más utilizados y nos proporciona la ventaja de tener un número de categorías emocionales

²³ <http://moviendoeemociones.blogspot.com/2014/02/clasificacion-teorica-general-de-las.html>

determinado, para elegir a cuál de ellas pertenece cada oración. Se utilizarán cinco categorías: enfado, alegría, tristeza, sorpresa y neutro. Esta última se utilizará en el caso de no detectar ninguna de las anteriores o cuando varias emociones se vean representadas con el mismo grado de intensidad.

3.2. Estudio de técnicas para el reconocimiento de emociones

Existen tres enfoques que se han utilizado a la hora de realizar tareas que traten de clasificación y reconocimiento de las emociones [7][3][19]: basado en lexicón, basado en machine learning y por último un enfoque híbrido. Cada uno de ellos se explica en los siguientes apartados.

3.2.1. Basado en lexicón

Un enfoque basado en lexicón se basa en buscar la aparición de una serie de palabras clave en un texto del que no sabemos a qué categoría pertenece (no etiquetado) y gracias a ellas ser capaces de clasificarlo con una emoción determinada.

Dentro de esta técnica, el enfoque más utilizado es la búsqueda de palabras clave, que consiste en ayudarse de una lista de palabras representativas sobre un tema, grupo social o emociones como es el caso de los utilizados para este proyecto, denominados como lexicones. Una vez se tienen elegidos esos lexicones, se lleva a cabo normalmente un preprocesamiento del texto que se quiere clasificar, eliminando palabras de parado, aplicando un proceso de lematización y representando las palabras restantes como una bolsa de palabras o vectores de palabras.

El proceso para decidir a qué categoría pertenece, o en el caso de este trabajo, qué emoción representa, se basa en detectar las palabras clave elegidas en la bolsa de palabras que representa el texto a categorizar. Por último se analiza la intensidad de cada emoción y se determina la etiqueta de la emoción para el texto.

Para la realización de este proyecto se ha decidido utilizar un enfoque basado en lexicón, ayudándonos de un conjunto de recursos emocionales para el español, los cuales se explicarán en la subsección 3.3.

3.2.2. Basado en Machine Learning

Las técnicas basadas en Machine Learning tratan de implementar algoritmos que pueden crear modelos con capacidad de obtener predicciones a partir de una serie de variables de entrada. Dentro de este enfoque existen dos modos diferentes de desarrollo.

- Técnicas de aprendizaje supervisado: consta de dos etapas, en la primera, llamada etapa de entrenamiento, se ayuda de un conjunto de documentos etiquetados con el objetivo de aprender. El conjunto de documentos utilizado tiene que estar representado mediante una serie de características, que permitan generar un modelo estadístico que sirve para modelar la distribución de los datos. Posteriormente, en la segunda etapa, llamada etapa de test, el objetivo es etiquetar nuevos documentos ayudándose del aprendizaje conseguido en la primera etapa. Siempre que el nuevo documento sin etiquetar esté representado de la misma forma, el algoritmo creado en la primera etapa será capaz de clasificarlo obteniendo la distribución de datos que mejor se le ajusta. Ejemplos de este enfoque son la clasificación de Naive Bayes, Support Vector Machines o redes neuronales.
- Técnicas de aprendizaje no supervisado: su finalidad es buscar patrones o estructuras ocultas en un conjunto de datos no etiquetados para poder construir un modelo que permita clasificar nuevos documentos. La ventaja de este tipo de aprendizaje frente al supervisado es que no es necesario poseer una serie de documentos etiquetados, evitando el elevado coste de ese proceso.

3.2.3. Enfoque híbrido

Las técnicas basadas en un enfoque híbrido tratan de realizar una combinación de las dos estrategias explicadas anteriormente. La principal ventaja de este enfoque consiste en que podemos obtener mejores resultados mediante la utilización de las ventajas de unas técnicas y otras y de esta manera equilibrar las desventajas de las mismas.

3.3. Estudio y análisis de lexicones

Se denomina lexicón a un diccionario o listado de términos representativos de una lengua, región, materia, etc. En el ámbito del PLN, consiste en un conjunto de términos que representan, por ejemplo, sentimientos, emociones, etc. Los lexicones nos ayudan a categorizar corpus no etiquetados y dividirlos en función del tema que tratan, género del autor, sentimiento, emoción, etc. Los lexicones se pueden obtener de tres formas diferentes [4][18]:

- Método manual: se basa en ir añadiendo palabras de manera manual, formando diferentes listas de términos en función de categorías por ejemplo, en función de diversas emociones.
- Método basado en diccionario: se basa en extraer un conjunto de palabras con orientación conocida (semillas) e ir incrementando el número de palabras a partir de algún diccionario o base de conocimiento léxico, teniendo en cuenta por ejemplo los antónimos (con significado contrario a la semilla) y sinónimos (con el mismo significado al de la semilla).
- Método basado en corpus: a partir de una lista de palabras se intenta buscar otras relacionadas con ellas, en un corpus de dominio específico.

Los lexicones marcados con emociones en español son escasos, ya que la mayoría se han construido para el inglés. Los lexicones que se han estudiado para la realización del presente trabajo son los explicados a continuación.

3.3.1. EmoSenticNet

EmoSenticNet [16] se trata de un lexicón formado por 13.188 términos. Los términos están etiquetados con una de las siguientes seis emociones: alegría, asco, enfado, miedo, sorpresa y tristeza. Para su creación se basa en los términos contenidos en SenticNet²⁴ (lexicón con información sobre la polaridad de cada una de sus entradas) y utiliza técnicas de Machine Learning supervisado (se le realiza un preprocesamiento de lematización). Para la etiquetación de cada término se planteó que las palabras con similares características suelen tener significado y uso parecido, por lo que representan emociones parecidas. Las características que se utilizaron

²⁴ <https://sentic.net/downloads/>

eran, o bien basadas en parámetros que provee ISEAR²⁵ (dataset creado a partir de encuestas) como información sobre el encuestado, emoción que sintió, comportamiento, etc., o basadas en diferentes medidas de similitud entre conceptos.

3.3.2. LIWC

Linguistic Inquiry and Word Count²⁶ (LIWC) [15] es una potente herramienta de investigación basada en una ciencia sólida. Su funcionamiento se basa en leer los textos que se le proporcionen y contar el porcentaje de palabras que expresen emoción, preocupaciones sociales e incluso partes de la oración.

Fue creado por investigadores interesados en psicología social, de salud, etc., por lo que las categorías del lenguaje fueron creadas para capturar estados psicológicos y sociales de las personas.

LIWC está basado en un conjunto de diccionarios, entre ellos se encuentra un diccionario de palabras sociales (familia, amigos, etc.), signos de puntuación, palabras afectivas (positivas y negativas), etc. Después del proceso de identificar qué palabras se encuentran en el texto de entrada y también en sus diccionarios, calcula los porcentajes con respecto al total de palabras, para relacionar el texto con una categoría.

LIWC proporciona también un lexicón de términos etiquetados con una serie de factores, entre ellos se encuentran dos de las emociones que se van a reconocer en este trabajo (enojo y tristeza).

3.3.3. NRC Emotion lexicon (Emolex)

Emolex²⁷ [1] se trata de un lexicón en inglés formado por 10.000 términos. Cada entrada incluye un término, una emoción y una medida de cuán fuertemente el término está asociado con la emoción.

²⁵ <https://www.affective-sciences.org/research/materials-and-online-research/research-material/>

²⁶ <http://liwc.wpengine.com/how-it-works/>

²⁷ <https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

Las emociones que trata son alegría, tristeza, ira, miedo, confianza, asco, sorpresa y anticipación, modelo de Plutchik, explicado anteriormente. Se incluyen los seis de Ekman, así como la confianza y la anticipación.

El primer paso para construir el lexicon fue elegir la lista de palabras y frases que iban a ser anotadas. Para ello, se utilizaron un total de 10.170 términos que se obtuvieron de tres fuentes diferentes:

- Del diccionario Macquarie Thesaurus²⁸ se obtuvieron los 200 unigramas y bigramas de cada categoría gramatical (verbos, adverbios, adjetivos y nombres) que más aparecen en el corpus de n-gramas de Google.
- Del lexicon WordNet Affect se escogieron todos los términos que tuvieran como máximo dos emociones asociadas.
- Del General Inquirer²⁹ [17] se utilizaron todos los términos que no fueran muy ambiguos (como máximo 3 emociones).

Para la anotación de cada término con la emoción que expresa se utilizó Amazon's Mechanical Turk, que se trata de una herramienta de crowdsourcing mediante la que los creadores de este lexicon subieron diferentes pequeñas tareas para la etiquetación de cada término y anotadores anónimos las llevaban a cabo. Cada término fue anotado por varias personas. En cada tarea se realizaba la anotación de un término que consistía en rellenar una encuesta, con 4 respuestas por pregunta y una única elección. La primera pregunta estaba formada por cuatro palabras, una de ellas era un sinónimo de la palabra a etiquetar. De esta forma se eliminaban gran parte de los anotadores aleatorios, ya que con esta simple pregunta, en caso de contestarse de forma errónea significaba que con un 75% de probabilidad había sido contestada aleatoriamente.

Para la validación de las anotaciones se crearon scripts automáticos. Algunos de los motivos por los que se descartaron anotaciones del total fueron:

- Alguna pregunta de la encuesta estaba sin contestar.
- Error en la primera pregunta de la encuesta.

²⁸ <https://www.macquariedictionary.com.au>

²⁹ <http://www.wjh.harvard.edu/~inquirer/>

- No se tenían en cuenta las anotaciones de anotadores que se equivocaron en una de cada tres preguntas.

Cada término podía ser asignado con un grado diferente de intensidad a varias emociones, para asignar la emoción se elegía la que más porcentaje de coincidencia tuviera.

Este lexicón se realizó en inglés como ya se ha dicho, aunque se encuentra disponible para más de 100 idiomas diferentes. Para la realización de este trabajo se va a utilizar su versión en español.

3.3.4. SEL

SEL³⁰ [2] se trata de un diccionario formado por 2.036 palabras en español etiquetadas con seis emociones básicas: alegría, enojo, miedo, tristeza, sorpresa y repulsión, siguiendo el enfoque de Ekman. El proceso que se siguió para formar este lexicón fue el siguiente:

1. Se tomó el recurso WordNet-Affect (Strapparava y Valitutti, 2004) construido para el inglés y etiquetado por categorías emocionales bajo el enfoque de Ekman.
2. Se tradujeron todas las palabras al español utilizando tres herramientas: Google Traductor, English-Spanish Interpreter Pro y el traductor de Babylon. Los resultados de estos traductores se unieron manteniendo las categorías emocionales y eliminado duplicados.
3. Se descartaron manualmente las palabras que una vez traducidas no se correspondían con la emoción asociada.
4. El siguiente paso consistió en una evaluación para confirmar que las palabras que formaban el recurso tenían relación con la emoción a la que estaban asociadas. Para ello se ayudaron de 19 evaluadores voluntarios que calificaron la relación de la palabra con la emoción asociada entre cuatro opciones: nula, baja, media y alta. Se implementaron varios umbrales de acuerdo para denotar la coincidencia entre los evaluadores.

³⁰ <http://www.cic.ipn.mx/~sidorov/>

5. Por último se creó un factor de probabilidad de uso afectivo (FPA) que indica la probabilidad de que una palabra sea usada para expresar una emoción con la que está etiquetada. A mayor FPA mayor probabilidad.

Finalmente, los tres lexicones que se van a utilizar para realizar pruebas con el sistema van a ser: Emolex, LIWC y SEL, ya que proporcionan una versión en español, que es el idioma elegido para el desarrollo del trabajo y se ajustan a las necesidades para la realización del mismo, proporcionando términos para al menos 2 de las 4 emociones que se pretenden reconocer. Posteriormente, en el desarrollo del proyecto se explicarán las pruebas realizadas y los lexicones elegidos para que formen parte del sistema final.

Capítulo 4. Desarrollo del proyecto

En este capítulo se explica el desarrollo llevado hasta la finalización de este proyecto. Se describe la metodología llevada a cabo, las historias de usuario que forman el proyecto, la solución elegida junto con su justificación y por último las herramientas que se han utilizado.

4.1. Metodología para el desarrollo de software y método a seguir

4.1.1. Metodología

En el marco de la ingeniería del software, la metodología para el desarrollo de proyectos es un proceso cuyo objetivo es estructurar, planificar y controlar el desarrollo de los mismos con el fin de llegar a una solución final.

Existen diversos tipos de metodologías, aunque las dos más utilizadas son la metodología tradicional y la ágil, las cuales se explican a continuación.

La metodología tradicional se caracteriza por:

- Su ciclo de vida sigue un desarrollo de software en cascada, en el que las diversas etapas se realizan de manera secuencial.
- No se puede comenzar una tarea si no se ha completado la anterior.
- Una vez finalizada una etapa no puede volver a modificarse, por tanto, cada etapa solo se ejecuta una vez.

Estas características conllevan una serie de inconvenientes para este tipo de metodología:

- No podremos volver atrás para modificar una tarea finalizada si nos damos cuenta de algún error.
- El usuario no verá el producto hasta que no haya sido finalizado, lo que provoca que cuando se finalice el usuario tenga cosas que no quiere o que estén mal respecto a sus requisitos.

En la siguiente ilustración (Ilustración 4.1.), se muestra el ciclo de vida de una metodología tradicional.

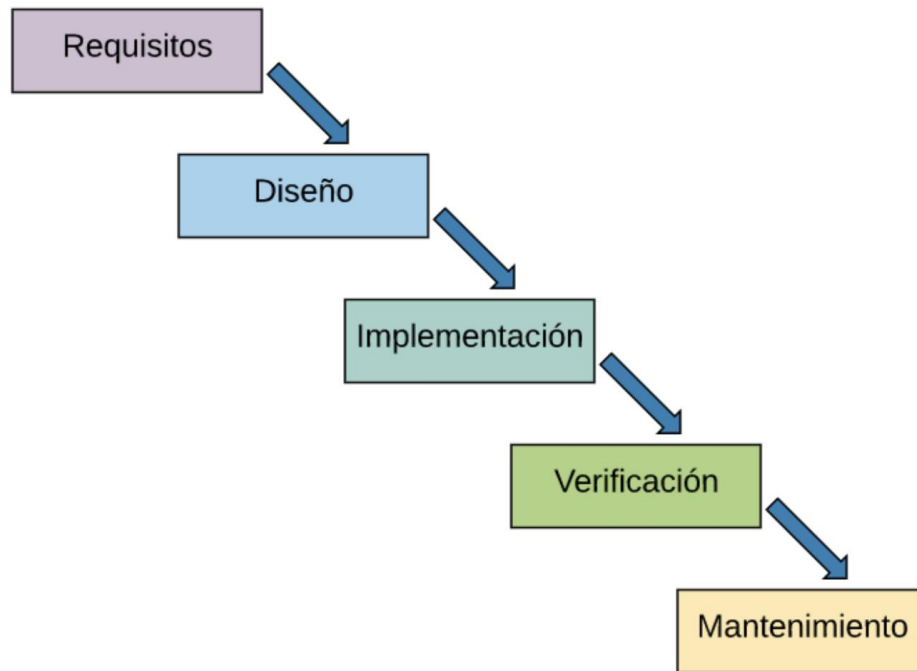


Ilustración 4.1. Ciclo de vida de una metodología tradicional³¹

Por otro lado se encuentra la metodología ágil, que se caracteriza por:

- Su ciclo de vida sigue un desarrollo de software iterativo e incremental, en el que el proyecto se divide en diversos bloques temporales que se denominan iteraciones.
- Las etapas de cada ciclo se repiten y se va añadiendo funcionalidad al sistema. Al finalizar cada etapa se muestra al cliente para que vaya observando el desarrollo del producto.
- Cada iteración es corta, de unas 2-4 semanas.
- Al finalizar cada iteración se hace una entrega parcial del proyecto al usuario para que vaya validando que se cumplen sus requisitos. Cada una de estas entregas se trata de un resultado usable.
- Se pueden solapar varias etapas.
- Los requisitos pueden ir cambiando durante el desarrollo del proyecto.

³¹ <https://openclassrooms.com/en/courses/4309151-gestiona-tu-proyecto-de-desarrollo/4538221-en-que-consiste-el-modelo-en-cascada>

En la siguiente Ilustración (Ilustración 4.2.) se puede observar el ciclo de vida de una metodología ágil.



Ilustración 4.2. Ciclo de vida de una metodología ágil³²

Una vez se han examinado las dos metodologías se ha decidido utilizar para el desarrollo de este proyecto la metodología ágil basada en un ciclo de vida iterativo e incremental, por los siguientes motivos:

- Este proyecto está formado por varias iteraciones y tendrá diversas entregas parciales.
- Una vez entregada cada una de las iteraciones, pueden surgir cambios debidos a un mal planteamiento de requisitos. La metodología escogida hace que no haya problemas en realizar estos cambios.
- Al finalizar cada iteración el tutor del proyecto (con el rol de cliente) podrá ver una parte usable del proyecto para verificar que se van cumpliendo los requisitos.

4.1.2. Método ágil

Existen multitud de métodos ágiles, como se muestra en la Ilustración 4.3. De entre ellos, se ha elegido **KANBAN**³³ como método ágil para el desarrollo del proyecto.

³² <https://www.tithink.com/es/2018/10/16/metodologias-agiles-que-son-y-para-que-sirven/>

³³ <https://uvadoc.uva.es/bitstream/handle/10324/1495/TFG-B.117.pdf;jsessionid=7ED080D34E72D98DCB793E243AEEA8F0?sequence=1>

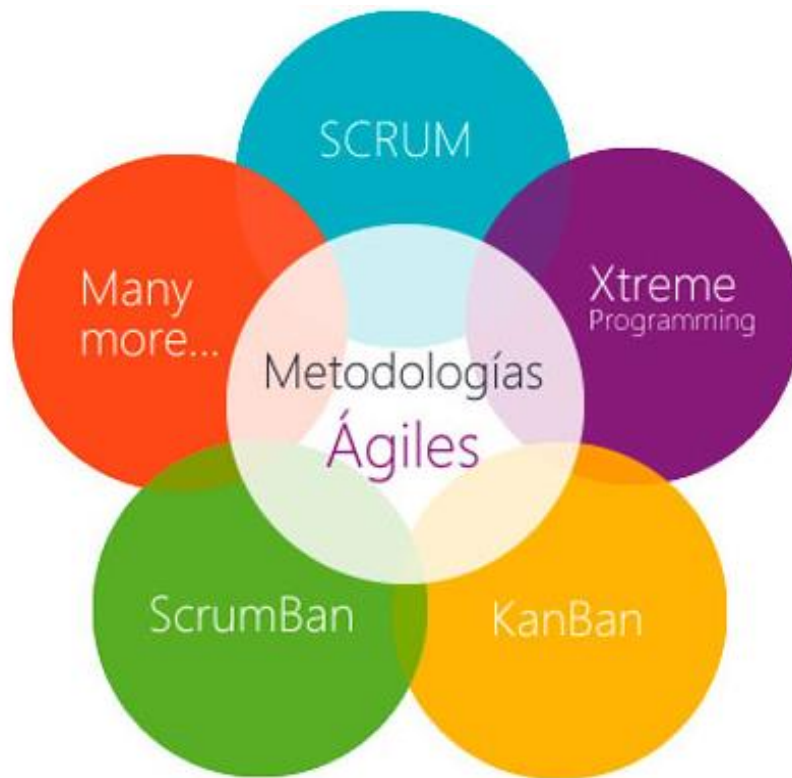


Ilustración 4.3. Método ágiles³⁴

El objetivo principal del método ágil Kanban se trata de gestionar de una manera general cómo se van completando las tareas que forman un proyecto. A continuación se explican las principales características que definen este método.

Kanban permite visualizar el trabajo y las diferentes fases del flujo de trabajo. Sigue un desarrollo incremental, dividiendo el trabajo en diversas tareas, y utilizando una técnica visual para ver la situación de cada una de ellas. Cada tarea se escribe en una tarjeta o post-it y se pega en una pizarra, como se muestra en la siguiente ilustración (Ilustración 4.4.).

³⁴ <https://www.ayselucus.es/noticia/métodos-ágiles-para-la-gestión-de-proyectos-i>

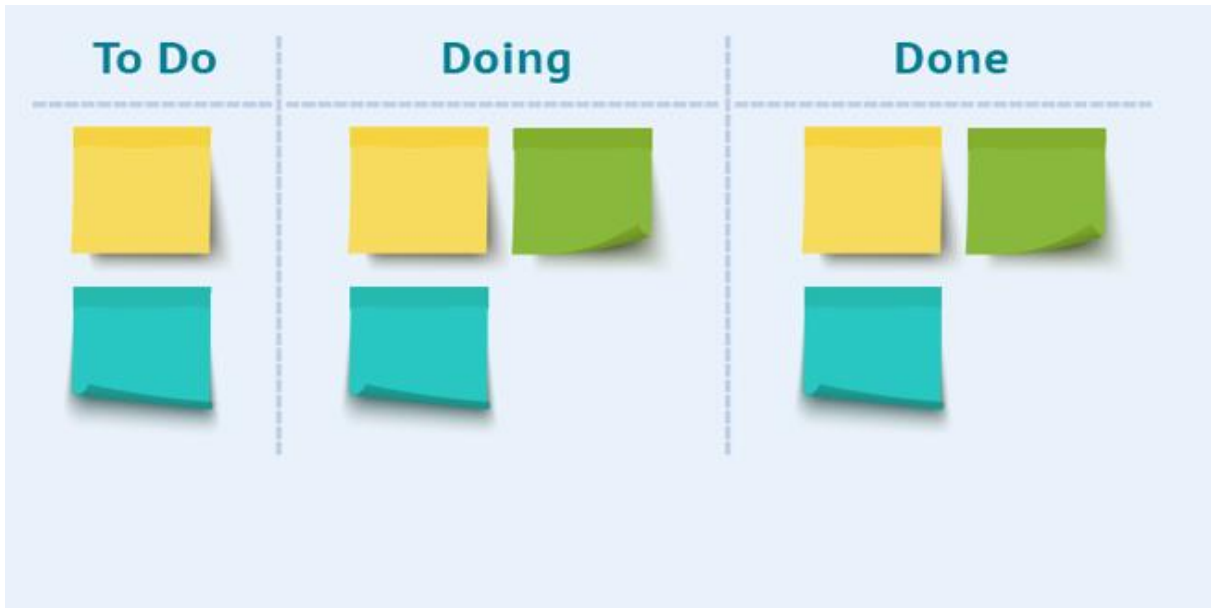


Ilustración 4.4. Ejemplo muro Kanban

Cada post-it tiene diversa información, aunque como mínimo debe tener una descripción de la tarea y una estimación de la duración de la misma. La pizarra donde se pegan estos post-it tiene tantas columnas como estados en los que se pueda encontrar una tarea. Esta visualización de las tareas permite de forma clara el trabajo que tiene cada persona y la prioridad de cada tarea.

Por otro lado, Kanban defiende la idea de que el trabajo en curso (WIP -Work In Progress) debe ser limitado, para que el número de tareas que se realicen como máximo en cada fase del ciclo de trabajo sea conocido. Además, para empezar una nueva tarea se debe haber finalizado otra tarea previa.

Kanban permite medir el tiempo que se tarda en completar una tarea (Lead time). Este tiempo se cuenta desde que se hace una petición hasta que se realiza la entrega. Por otro lado otra métrica importante se trata del 'cycle time' que mide el tiempo desde que una tarea comienza hasta que termina. Con esta medida se obtiene el rendimiento del proceso.

Kanban es muy flexible ya que no requiere de reuniones diarias con el cliente (papel que realizará el tutor del proyecto).

Por último, se muestra en la siguiente ilustración (Ilustración 4.5) el muro utilizado durante el desarrollo del proyecto³⁵ en su estado inicial. En el se pueden observar las diversas historias de usuario junto con el desarrollo de la memoria como tareas. Cada una de ellas deberá pasar por los estados “To-do” (por hacer), “In progress” (en progreso) y “done” (hecho), de forma que una vez finalice el proyecto todas deberán estar en el estado “Done”.

To-do +	Do today +	In progress 0 / 3 +	Done +
Procesado de lexicones			
Implementación front-end			
Implementación de clasificador de emociones (hack-end)			
Implementar capacidad para dar diferentes respuestas			
Evaluación del sistema			
Redactar memoria			

Ilustración 4.5. Muro de kanban utilizado

4.2. Historias de usuario

Las historias de usuario son una de las principales partes de la metodología ágil. Mediante ellas se obtienen los requisitos que forman un proyecto software.

Una historia de usuario es una especificación de una parte del software. En ella se describe la funcionalidad que podría ser útil para el usuario del sistema. A partir de ellas se puede empezar a trabajar, y pueden ser modificadas y adaptadas a nuevos requisitos durante el desarrollo del proyecto.

Para crear una historia de usuario es necesario que exista una comunicación y un acuerdo entre las personas que forman el equipo y el cliente. Suelen ser de pequeño tamaño y centrarse en los principales aspectos.

Una historia de usuario puede tener los siguientes componentes:

³⁵ <https://kanbanflow.com/>

- **ID:** identificador de la historia de usuario.
- **Título:** título de la historia de usuario. Debe ser lo más claro posible.
- **Prioridad:** importancia de la historia de usuario respecto al resto. Puede tomar los siguientes valores: baja, media, alta y crítica.
- **Descripción:** breve descripción de la finalidad de la historia de usuario.
- **Dependencias:** esta se produce en el caso en que una historia de usuario dependa de otra para su desarrollo.
- **Criterios de aceptación:** son las pruebas que se le realizan a una historia de usuario para comprobar que se ha desarrollado correctamente.

En la siguiente tabla (Tabla 4.1), se muestra la estructura con la que mostraremos las historias de usuario.

ID	TÍTULO
DESCRIPCIÓN	
PRIORIDAD	DEPENDENCIAS

Tabla 4.1. Plantilla historia de usuario

A continuación se muestran las distintas historias de usuario del proyecto que se va a realizar.

1	PROCESADO DE LEXICONES
Como usuario quiero disponer de varios lexicones con un formato adecuado para utilizar en el clasificador de emociones	
Media	-

Tabla 4.2. Historia de usuario: “Procesado de lexicones”

2 IMPLEMENTACIÓN FRONT-END	
Como usuario quiero disponer de un front-end en la skill para poder comunicarme con la persona que la utilice	
Alta	-

Tabla 4.3. Historia de usuario: “Implementación front-end”

3 IMPLEMENTACIÓN DE CLASIFICADOR DE EMOCIONES (BACK-END)	
Como usuario quiero disponer de un clasificador de emociones integrado en la skill	
Alta	1

Tabla 4.4. Historia de usuario: “Implementación de clasificador de emociones (back-end)”

4 CAPACIDAD PARA DAR DIFERENTES RESPUESTAS	
Como usuario quiero poder dar diferentes respuestas para cada una de las emociones que se reconocen	
Baja	-

Tabla 4.5. Historia de usuario: “Capacidad para dar diferentes respuestas”

5 EVALUACIÓN DEL SISTEMA	
Como usuario quiero comprobar el funcionamiento del reconocedor de emociones con los diversos lexicones de prueba	
Media	2,3

Tabla 4.6. Historia de usuario: “Evaluación del sistema”

4.3. Propuesta de solución

En este apartado se explica la propuesta de solución elegida para este TFM. Consiste en desarrollar una Skill para el AVI Amazon Alexa que incorpore un reconocedor de emociones para contestar al usuario de una forma u otra en función de la emoción que exprese al comunicarse con el sistema. Para ello se necesitan desarrollar los siguientes módulos:

- Sistema para entablar comunicación entre el usuario y la Skill, para obtener la oración que se quiere analizar y devolver una respuesta.
- Sistema para la clasificación de emociones, a partir de una oración será capaz de obtener la emoción que esté presente en ella con más intensidad.

4.4. Descripción de la solución

En este apartado se explica el desarrollo de la solución elegida mediante varias iteraciones. Cada iteración representa una historia de usuario de las comentadas anteriormente (Apartado 4.2), por tanto habrá 5 iteraciones. Antes de comenzar con la explicación de estas iteraciones se muestra en la Ilustración 4.6 un diagrama de las partes que forman el sistema y la interacción entre ellas.

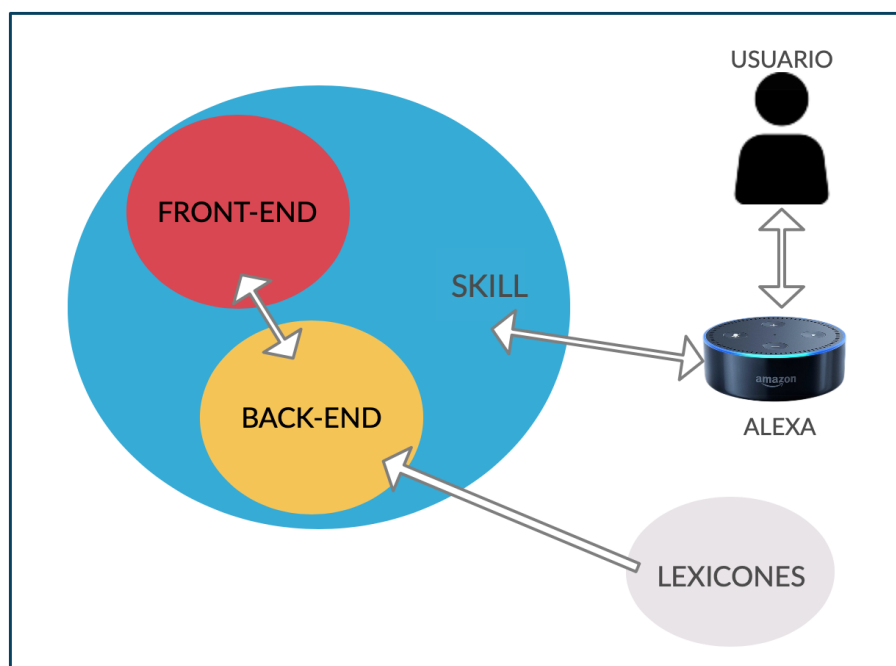


Ilustración 4.6. Diagrama del sistema

4.4.1. Iteración 1

Iteración 1: “Como usuario quiero disponer de varios lexicones con un formato adecuado para utilizar en el clasificador de emociones”.

La iteración 1 se corresponde con la historia de usuario “Procesado de lexicones”, que se muestra en la Tabla 4.2.

Como ya se ha explicado anteriormente (Apartado 3.2.1) para el sistema de clasificación de emociones se va a utilizar un enfoque basado en lexicones. Es por ello que el principal recurso que se necesite sea uno o varios lexicones de palabras etiquetadas con las emociones que se pretenden reconocer. Para ello se llevó a cabo un primer paso que consistió en un estudio de diversos lexicones (Apartado 3.3). Finalmente se decidió realizar pruebas con tres lexicones, los cuales están en español y contienen al menos dos de las cuatro emociones que se requieren. Estos lexicones son LIWC que contiene términos para tristeza y enojo, SEL y Emolex que contienen términos para las cuatro emociones.

Cada uno de estos lexicones, cuyo recurso de donde se han descargado se referenció en el apartado 3.3, tiene una estructura diferente. Es por ello que se ha tenido que llevar a cabo un preprocesamiento para cada uno de ellos para proporcionarle la estructura que más nos beneficie a la hora de crear el sistema y cargar los lexicones en memoria.

A continuación se va a explicar el preprocesamiento realizado para cada uno de los lexicones y el formato que se le ha dado finalmente. Todos ellos se han almacenado en un fichero en formato CSV separando los diferentes elementos que forman los registros mediante ‘;’.

LIWC contiene términos para tristeza y enojo además de más tipos de categorías como adverbios, conjunciones, verbos, etc. Cada una de las categorías tiene asociado un número. Para quedarnos únicamente con los términos necesarios había que buscar para cada término si pertenecía a la categoría de palabras de enojo o de tristeza, números 129 y 130 respectivamente. Por otro lado existen términos que contienen un * al final indicando que puede tener varias terminaciones (masculino/femenino, singular, plural, etc.). Para resolver este problema se hizo un

proceso manual añadiendo los diversos términos para eliminar el asterisco. En la Ilustración 4.7. se muestra un ejemplo del formato explicado.

trágic* 125 127 130
vergonzos* 125 127 128 129

Ilustración 4.7. Formato inicial LIWC

El formato final que se le ha dado a este lexicón es el que se muestra en la siguiente ilustración (Ilustración 4.8.). Como se puede observar cada registro está formado por dos elementos, primero el término y después la emoción a la que está asociado.

discuto;Enojo
disputa;Enojo
disuadir;Tristeza
doler;Tristeza
doliente;Tristeza
domino;Enojo

Ilustración 4.8. Formato final LIWC

Emolex proporciona su lexicón en muchos idiomas diferentes en un fichero excel. Dentro de este fichero están todos los términos en los diversos idiomas, y para indicar la/s emoción/es a la/s que pertenecen lo indican con un 1, como se muestra en la siguiente ilustración (Ilustración 4.9.).

Positive	Negative	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise	Trust
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1
0	1	0	0	0	1	0	1	0	0
0	1	1	0	0	1	0	1	0	0

Ilustración 4.9. Formato inicial Emolex

Como se puede observar proporciona más emociones de las necesarias, por lo que se tuvo que realizar un procesado del lexicón para quedarnos con los términos en español y las emociones necesarias. El formato final es el que se muestra en la Ilustración 4.10., donde tenemos el término y un 1 en la emoción o emociones a las que pertenece. El orden de las emociones es: enojo, alegría, tristeza y sorpresa.

```
obituario;0;0;1;1
obituario;0;0;1;0
objecion;1;0;0;0
```

Ilustración 4.10. Formato final Emolex

SEL contiene términos etiquetados con la emoción a la que pertenecen y un factor que indica la probabilidad de que ese término se utilice para expresar la emoción asociada, como se muestra en la siguiente ilustración (Ilustración 4.11). Este lexicón contiene más emociones aparte de las que se necesitan, por lo que se tuvo que llevar un proceso de búsqueda de los términos para alegría, tristeza, enojo y sorpresa, eliminando el resto.

Palabra	PFA	Categoria
abundancia	0.83	Alegría
acabalar	0.396	Alegría
acallar	0.198	Alegría
acatar	0.198	Alegría
accion	0.397	Alegría

Ilustración 4.11. Formato inicial SEL

El formato final que se le ha dado es el que se muestra en la Ilustración 4.12., en el que cada registro está formado por los tres elementos citados (Término;PFA;Emoción).

```
envidioso;0.462;Enojo
escandalizar;0.297;Enojo
escandalo;0.198;Enojo
estafar;0.43;Enojo
```

Ilustración 4.12. Formato final SEL

4.4.2. Iteración 2

Iteración 2: “Como usuario quiero disponer de un front-end en la skill para poder comunicarme con la persona que la utilice”.

La iteración 2 se corresponde con la historia de usuario “Implementación Front-end”, que se muestra en la Tabla 4.3.

En esta iteración se lleva a cabo el desarrollo de front-end de la skill, parte encargada de la interacción con el usuario, que deberá entender lo que el usuario le diga al comunicarse con ella.

El front-end de una skill de Amazon Alexa está formado por diversos componentes:

- **Invocation Name:** será el nombre con el que se invocará la skill por parte del usuario. Debe estar formado por 2 o más palabras. El nombre elegido para la skill es ‘yo te escucho’.
- **Intents:** son las diferentes acciones que proporciona la skill para que el usuario las utilice. La skill tiene implementados por defecto cuatro intents: ayuda, cancelar, parar y navegar hasta la raíz (se encarga de llevar al usuario hasta la portada del dispositivo). Para nuestra skill, se ha creado un único intent, que implementará la habilidad de reconocer emociones.
- **Slots:** son registros dentro de utterances con los que se pretende obtener información del usuario. Algunos slots deben de ser rellenados obligatoriamente para continuar la ejecución de la skill. En nuestra skill, existe un único slot, que será el mensaje que proporcione el usuario para analizar y reconocer la emoción. Este será obligatorio.
- **Utterances:** son sentencias definidas para cada intent y que sirven para lanzarlos. En el caso de nuestra skill, se ha definido un único utterance que será el slot del mensaje que proporciona el usuario.

Una vez se han explicado los elementos que forman el front-end de nuestra skill, se explicará la interacción compuesta para este sistema.

1. El usuario abrirá la skill haciendo uso del nombre de invocación explicado previamente. A través de él se iniciará la skill, la cual mandará un mensaje de bienvenida y hará una pregunta para obtener una respuesta acerca del estado de ánimo del usuario. La pregunta que hará la skill será una de las que se muestran a continuación:

- ¿cómo estás?
 - ¿qué tal?
 - ¿cómo te ha ido el día?
 - ¿cómo te fue el día?
 - ¿cómo fue el día?
 - ¿qué tal estás?
 - ¿qué tal el día?
 - ¿qué tal fue el día?
 - ¿estás bien?
- ¿cómo te encuentras?
 - ¿cómo vas?
 - ¿cómo va el día?
 - ¿cómo llevas el día?
 - ¿cómo fue el día de ayer?
 - ¿qué tal el día de ayer?
 - ¿cómo estás hoy?
 - ¿qué tal hoy?
 - ¿cómo te va?
2. El usuario responderá con una oración, el cual hará que se rellene el slot del mensaje que existe en la skill y se inicie directamente el intent del reconocimiento de emociones.
 3. Si no se obtiene un mensaje, se volverá a preguntar al igual que en el punto 1, para obtener un mensaje válido.
 4. Una vez obtenemos el mensaje, y es procesado, la skill proporcionará una respuesta al usuario y preguntará si quiere decir algo más. En caso de reconocer un estado neutro, se volverá a realizar otra pregunta para obtener otro mensaje para analizar. La pregunta que realizará la skill en la segunda interacción será una de las que se enumeran a continuación.
 - ¿Quieres contarme otra cosa?
 - ¿Tienes algo más que decirme?
 - ¿Algo más que contarme?
 - ¿Algo más?

A estas preguntas, el usuario podrá contestar diciendo que no quiere decir nada más, finalizando así la ejecución de la skill.

4.4.3. Iteración 3

Iteración 3: “Como usuario quiero disponer de un clasificador de emociones integrado en la skill”.

La iteración 3 se corresponde con la historia de usuario “Implementación de clasificador de emociones (back-end)”, que se muestra en la Tabla 4.4.

En esta iteración, se lleva a cabo el desarrollo del sistema de clasificación de emociones. Como ya se ha explicado anteriormente en diferentes apartados, para esta clasificación se va a utilizar un método basado en lexicones, utilizando los lexicones explicados en el apartado 4.4.1.

Este método, explicado en el apartado 3.2.1 está formado por varios pasos, que se explican a continuación:

1. Tomar la oración transmitida por el usuario hacia nuestro sistema, ya que hay tendremos que analizarla, como un vector o bolsa de palabras. Esto quiere decir que cada palabra será tenida en cuenta de forma individual, teniendo por tanto, tantos términos en la bolsa o vector como palabras formen la oración.
2. Buscar coincidencias de las palabras que forman la oración en los términos que forman los lexicones utilizados. Se lleva un contador para cada emoción, al que se irá sumando el peso de cada término que represente dicha emoción y que esté presente tanto en la oración como en el lexicon. Un término será clasificado en la emoción para la que obtenga una mayor puntuación.
3. Por último, la emoción elegida como la más representativa de la oración y por tanto, la que está expresando el usuario, será la que tenga el mayor valor en el contador explicado en el punto anterior.

Por otro lado, se tomará como resultado de la clasificación un estado neutro cuando no se haya encontrado ninguna coincidencia entre los términos de la oración y los lexicones o en el caso de que el valor del contador de varias emociones sea el mismo.

Por último, se presentan varios ejemplos del proceso explicado. En el primero de ellos, Ilustración 4.13., se muestra una oración que expresa una emoción neutra. Como se puede observar el contador de todas las emociones es 0, lo que indica que no se ha encontrado ninguna coincidencia en los lexicones utilizados.

```
Oración del usuario: Hoy me he comprado un sandwich para comer  
Contador por emociones: {'Enojo': 0, 'Alegria': 0, 'Tristeza': 0, 'Sorpresa': 0}  
Resultado: estado neutro
```

Ilustración 4.13. Ejemplo ejecución resultado estado neutro

El segundo ejemplo, Ilustración 4.14., muestra una oración que obtiene una emoción concreta, en este caso alegría. Como se puede observar el contador de la emoción con el valor más alto es el elegido.

```
Oración del usuario: Acabo de aprobar el carnet del coche  
Contador por emociones: {'Enojo': 0, 'Alegria': 0.664, 'Tristeza': 0, 'Sorpresa': 0}  
Resultado: emoción reconocida -> Alegría
```

Ilustración 4.14. Ejemplo ejecución resultado emoción alegría

4.4.4. Iteración 4

Iteración 4: “Como usuario quiero poder dar diferentes respuestas para cada una de las emociones que se reconocen”.

La iteración 4 se corresponde con la historia de usuario “Capacidad para dar diferentes respuestas”, que se muestra en la Tabla 4.5.

Una vez que hemos cumplido las anteriores iteraciones, ya hemos obtenido una oración y la hemos procesado consiguiendo reconocer la emoción que se muestra con más intensidad.

El siguiente paso, que es el que se pretende cumplir mediante el desarrollo de esta iteración consiste en dar una respuesta al usuario. Ésta variará en función de la emoción que se haya reconocido, aportando un mensaje que ayude al usuario, le informe acerca de la emoción que siente o haga que la skill empatee con él. A continuación se muestra la intencionalidad de la respuesta que se dará dependiendo de la emoción:

- Alegría: se reforzará esa emoción ya que es positiva.
- Tristeza: se transmitirá un mensaje que ayude a animar al usuario.
- Enojo: se transmitirá un mensaje que ayude a calmar al usuario.
- Sorpresa: se transmitirá un consejo o un mensaje que ayude al usuario a relajarse.

- Neutro: en caso de determinar que no existe ninguna emoción predominante, se intentará que el usuario nos diga más cosas para poder reconocer una emoción.

Una vez se tenía definida la intencionalidad que debía tener la respuesta para cada emoción, se pretendía reforzar la skill para hacerla más amplia e imprevisible. Para ello se decidió añadir diferentes respuestas para cada emoción, para que el sistema no estuviera limitado a únicamente 5 respuestas, una para cada emoción una de las cuatro emociones y otra para el estado neutro. Se elaboró una batería de respuestas como se muestra en la siguiente tabla (Tabla 4.7.).

ALEGRÍA	TRISTEZA	SORPRESA	ENOJO
¡Me alegro de que estés contento!	Me gustaría tener brazos para abrazarte	Caray, espero que tu sorpresa sea por algo bueno	Vaya, siento que estés enfadado
Yuhuu	Jo, siento oír eso	Me sabe mal que estés sorprendido	Siento oír eso
¡Genial!	Vaya, comprendo que estés fastidiado	¿Sabías que sorprenderte activa el sistema nervioso autónomo?	Una buena idea sería dar un paseo
Yupi	Anímate, seguro que tiene solución	¿Sabías que sorprenderte activa el sistema somático?	Respiración profunda
No tengo manos pero, ¡qué demonios!, choca esos cinco	Siento que estés triste, hablar con un amigo podría ayudarte	Apuesto a que en este momento tienes las cejas elevadas o la boca abierta	¿Sabes que enfadarse roba un montón de energía? Sería mejor utilizar esa energía en cosas que te hagan feliz
Caramba, estás en racha	Poner una canción para subir el animo	La sorpresa suele preceder a otra emoción	Contar hasta 10
Poner una canción alegre	Contar un chiste	La sorpresa suele ser una emoción de duración breve	Quizás salir a dar un paseo te ayude

Tabla 4.7. Batería de respuestas

Dentro de estas respuestas, encontraremos alguna que simplemente refuerce esa emoción, en el caso de la alegría (yupi, genial, etc.), otras que aportarán información al usuario acerca de la emoción (¿Sabías que...?, la sorpresa suele ser una emoción de duración breve, etc.), otras que intenten empatizar con el usuario (siento oír eso, me gustaría poder abrazarte, etc.) y por último existen algunas que no son una respuesta simple, si no que llevan a cabo una acción, como poner una canción alegre³⁶, contar un chiste breve para que el usuario se anime, iniciar una respiración pausada (inspiración 4 segundos y expiración 4 segundos) o contar hasta 10 lentamente para calmar al usuario.

Por otro lado, si no se reconoce ninguna emoción, se tiene como resultado el estado neutro. En este caso el proceso seguido consiste en responder con otra pregunta (¿algo más?, ¿quieres contarme otra cosa?) para intentar que el usuario nos diga algo más que poder analizar.

Por último, para finalizar el desarrollo de esta iteración también se integró una habilidad especial que proporciona Amazon Alexa para los desarrolladores de skills. Se trata de ‘Speech Synthesis Markup Language’³⁷ (SSML), que es un lenguaje de marcado que permite a nuestra skill hablar con diferentes efectos, entonaciones y frases hechas. Esto nos permite darle un toque especial a las respuestas y hacer incluso que la skill parezca más cercana al usuario, ya que da la sensación de tener sentimientos. Las utilizadas son las que se proporcionan para el idioma español.

4.4.5. Iteración 5

Iteración 5: “Como usuario quiero comprobar el funcionamiento del reconocedor de emociones con los diversos lexicones de prueba”.

La iteración 5 se corresponde con la historia de usuario “Evaluación del sistema”, que se muestra en la Tabla 4.6.

³⁶ <http://psicoterapiaserendipia.com/levantar-el-animo-canciones/>

³⁷ <https://developer.amazon.com/es-ES/docs/alexa/custom-skills/speech-synthesis-markup-language-ssml-reference.html>

Una vez el sistema está totalmente desarrollado, se llevó a cabo una evaluación del sistema, para comprobar el funcionamiento, la interacción con el usuario y ver los resultados con diferentes combinaciones de lexicones.

Como ya hemos dicho en el apartado 4.4.1 se tenían 3 lexicones con un formato adecuado para su uso en la skill. Con ellos se podían hacer diferentes combinaciones y diversas pruebas para finalmente elegir los que mejores resultados dieran.

Para realizar esta evaluación se ha utilizado un dataset [20] de tweets etiquetados con una de las cuatro emociones utilizadas en este proyecto. Este dataset está formado por un total de 4025 tweets: 858 que expresan enojo, 1816 alegría, 1008 tristeza y 344 sorpresa. Cada registro fue categorizado por el sistema de clasificación de emociones creado, con el objetivo de que se obtuviera la misma emoción que indicaba el dataset. Los resultados que se obtuvieron con las diferentes combinaciones son los que se muestran a continuación (Tabla 4.8.).

COMBINACIÓN	ACIERTOS	NEUTROS	FALLOS	%ACIERTO
SEL	922	2484	619	59,83127839
EMOLEX	563	2961	501	52,91353383
SEL+EMOLEX	1009	2149	867	53,78464819
SEL+LIWC	1310	2046	669	66,195048
EMOLEX+LIWC	930	2520	575	61,79401993
SEL+EMOLEX+LIWC	1242	1936	847	59,45428435

Tabla 4.8. Resultados obtenidos

Los porcentajes de acierto y fallo que se muestran son teniendo en cuenta el total de aciertos y fallos, no neutros. Una vez obtenidos los resultados, se decidió elegir la combinación de los lexicones SEL y LIWC, en la que SEL sería el principal ya que tiene términos para todas las emociones y LIWC serviría para reforzar las emociones tristeza y enojo, ya que es la combinación con mejor porcentaje de acierto respecto al fallo, además de tener el mayor número de tweets en los que se acierta la emoción y de las que menos tweets etiquetados como neutros tienen.

Por último y para finalizar este apartado, se muestra en la Ilustración 4.15. un ejemplo real de una interacción entre el usuario y el sistema creado.

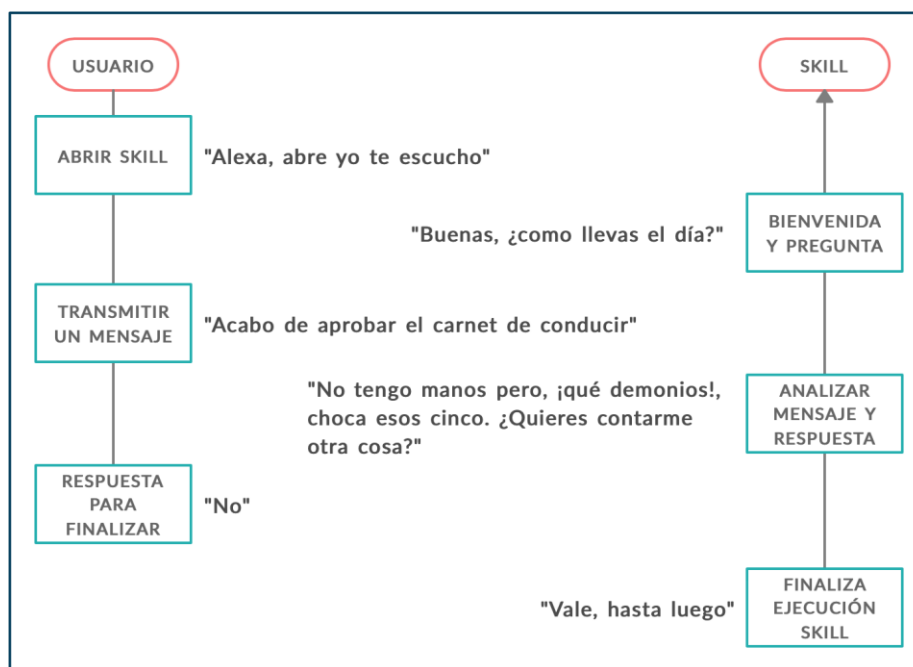


Ilustración 4.15. Ejemplo real de interacción del sistema final creado

4.5. Herramientas para la implementación del sistema

4.5.1. Alexa Skills Kit

Alexa Skills Kit³⁸ (ASK) es un conjunto de herramientas, documentación, muestras de código y API en self-service que permite desarrollar skills para Alexa de forma rápida y sencilla.

Las skills se pueden crear mediante dos formas diferentes:

- Creando un servicio propio y alojarlo en cualquier proveedor en la nube. Este servicio debe aceptar solicitudes HTTPS.
- Utilizando AWS Lambda, una oferta de Amazon Web Services, que se trata de un servicio que permite ejecutar código en la nube sin administrar ningún servidor. En este proyecto esta ha sido la opción elegida. Alexa envía sus solicitudes de usuario de código y su código puede inspeccionar la solicitud,

³⁸ <https://developer.amazon.com/es-ES/alexa/alexa-skills-kit>

realizas las tareas necesarias (como buscar información en línea) y luego responder. Es posible escribir funciones Lambda en varios lenguajes de programación: Java, Python, Node.js, C# o Go.

Para el desarrollo de este proyecto se ha utilizado la SDK para Python que proporciona ASK.

4.5.2. Python

El sistema se ha desarrollado con el lenguaje de programación Python (Ilustración 4.16.), en su versión 3.7.2 que puede ser utilizado en la plataforma explicada anteriormente (Apartado 4.5.1.). Este lenguaje fue creado por un investigador holandés llamado Guido Van que trabajaba en el centro de investigación CWI de Ámsterdam.



Ilustración 4.16. Lenguaje de programación Python³⁹

Python es un lenguaje de programación interpretado, de alto nivel pero fácil de aprender y de utilizar. Además es de código abierto por lo que su uso es gratuito. Permite la programación orientada a objetos y la programación estructurada.

Este lenguaje puede ser utilizado para desarrollo web, acceso a bases de datos, desarrollo software, etc. Debido a su sintaxis simple y su naturaleza interpretada, es ideal para realizar scripting sobre la mayor parte de las plataformas. También dispone de una línea de comandos a través de la que se pueden introducir y ejecutar sentencias, produciendo resultados.

³⁹ <https://www.linuxadictos.com/python-los-lenguajes-de-codigo-abierto.html>

Python puede ser utilizado con una serie de librerías que proporcionan una serie de utilidades. Para la creación de este sistema se han utilizado las siguientes:

- csv: utilizada para trabajar con ficheros '.csv'.
- operator: ofrece un conjunto de funciones eficientes que tienen su correspondencia en operadores comunes de Python. La hemos utilizado para acceder a datos organizados en listas.
- random: utilizada para obtener un número aleatorio dentro de un rango.
- logging: utilizada para la depuración del código.
- Diversas librerías proporcionadas por Amazon para la creación de la Skill.

Capítulo 5. Pruebas de validación y usabilidad.

5.1. Validación

Una vez finalizado el desarrollo de la skill, se llevó a cabo un proceso de validación de la misma, para comprobar manualmente que todo el funcionamiento tenía un resultado satisfactorio. Con él, se verificaron los siguientes puntos:

- Inicio de la skill.
- Detección por parte de la skill de la respuesta del usuario.
- Respuesta de la skill con 'sí' o 'no' si es posible.
- Detección de la emoción.
- Respuesta de la skill con mensaje relacionado a la emoción expresada.
- Cierre de la skill.
- Ver imagen que represente la emoción detectada.
- Varias interacciones consecutivas con resultado correcto.

5.2. Usabilidad

Las pruebas para comprobar la usabilidad de la skill se realizaron mediante la interacción de varios usuarios con ella y la posterior realización de una breve encuesta. En esta, se les pedirá que describan respondiendo a las preguntas la experiencia que han tenido durante el uso de la skill.

Esta encuesta se ha realizado a 3 personas de diferente edad y género (Tabla 5.1.), todas ellas inmediatamente después del uso de la skill.

USUARIO	EDAD	GÉNERO
1	24	Masculino
2	28	Femenino
3	46	Masculino

Tabla 5.1. Usuarios que realizaron la encuesta de usabilidad

En la Ilustración 5.1, se muestra la encuesta realizada al usuario 1.

Valore el acierto de las respuestas de la skill *

	1	2	3	4	5	
Muy poco acertado	☐	☐	☐	☒	☐	Muy acertado

Valore la coherencia de las respuestas de la skill *

	1	2	3	4	5	
Muy poca	☐	☐	☐	☐	☒	Mucha

Valore la interacción con la skill *

	1	2	3	4	5	
Muy difícil	☐	☐	☐	☐	☒	Muy fácil

¿Cuanto te ha gustado la skill? *

	1	2	3	4	5	
Muy poco	☐	☐	☐	☒	☐	Mucho

¿Te ha ayudado la skill a mejorar tu estado de ánimo? *

Sí, estaba triste y ayudo a alegrarme un poco.

Ilustración 5.1. Encuesta usuario 1.

En la Ilustración 5.2, se muestra la encuesta realizada al usuario 2.

Valore el acierto de las respuestas de la skill *

	1	2	3	4	5	
Muy poco acertado	☐	☐	☒	☐	☐	Muy acertado

Valore la coherencia de las respuestas de la skill *

	1	2	3	4	5	
Muy poca	☐	☐	☐	☐	☒	Mucha

Valore la interacción con la skill *

	1	2	3	4	5	
Muy difícil	☐	☐	☐	☒	☐	Muy fácil

¿Cuanto te ha gustado la skill? *

	1	2	3	4	5	
Muy poco	☐	☐	☐	☒	☐	Mucho

¿Te ha ayudado la skill a mejorar tu estado de ánimo? *

Sí, estaba feliz y me puso una canción que me alegro aún más!

Ilustración 5.2. Encuesta usuario 2.

En la Ilustración 5.3, se muestra la encuesta realizada al usuario 3.

Valore el acierto de las respuestas de la skill *

	1	2	3	4	5	
Muy poco acertado	☐	☐	☐	☐	☒	Muy acertado

Valore la coherencia de las respuestas de la skill *

	1	2	3	4	5	
Muy poca	☐	☐	☐	☐	☒	Mucha

Valore la interacción con la skill *

	1	2	3	4	5	
Muy difícil	☐	☐	☐	☐	☒	Muy fácil

¿Cuanto te ha gustado la skill? *

	1	2	3	4	5	
Muy poco	☐	☐	☐	☐	☒	Mucho

¿Te ha ayudado la skill a mejorar tu estado de ánimo? *

Sí, la respiración que me ayudo a realizar hizo que calmara mi ira.

Ilustración 5.3. Encuesta usuario 3.

Como resultado de las encuestas podemos hacernos la idea de que la skill es fácilmente utilizable por diversas personas, lo que quiere decir que las interacciones creadas son adecuadas para que el usuario se comunique de forma natural y sencilla con el sistema creado.

Capítulo 6. Conclusiones y trabajos futuros.

En este último capítulo se describen las conclusiones que se han obtenido una vez se ha finalizado el desarrollo del proyecto y las posibles mejoras que se pueden aplicar a este trabajo en el futuro.

6.1. Conclusiones

En este punto doy por finalizado el proyecto ya que se han conseguido alcanzar e incluso superar los objetivos que se habían marcado para llevar a cabo con la realización del mismo.

Decidí elegir este TFM debido a que tengo un gran interés por el campo del PLN. En el último curso del Grado en Ingeniería Informática ya dí una asignatura relacionada con este campo y mi TFG ya se enmarcaba dentro de este área, en concreto dentro de la detección del género del autor, en el que se utilizaron otras técnicas diferentes a las utilizadas en este proyecto. Esta fue otra de las razones, ya que quería aprender a realizar la tarea del reconocimiento y clasificación de emociones mediante el uso de otra técnica diferente. En este caso utilicé un método basados en lexicones.

Por otro lado, otro motivo que me llevó a elegir este proyecto fue el objetivo que se tenía de unificar este ámbito ya comentado del PLN a otra de las áreas más punteras dentro de la informática, como son los AVI. Con estos últimos solo había tenido un breve contacto durante el desarrollo del máster que pretendo finalizar con el desarrollo de este TFM, y es por ello que quería adentrarme y aprender cómo se desarrollan nuevas habilidades para estos dispositivos.

La realización de este proyecto pienso que ha sido bastante buena, ya que gracias a él ahora soy capaz de crear nuevas Skills para Alexa y además he aumentado mi conocimiento de PLN con una nueva técnica aprendida.

6.2. Trabajos futuros

En cuanto a posibles cambios o trabajos futuro que se podrían aplicar a este proyecto, cabe destacar que el sistema se ha creado para el idioma español, y una de las mejoras que se podría desarrollar es ampliar el sistema hacia otros idiomas como el inglés, italiano, francés, etc.

Además para el inglés la plataforma para desarrolladores de Alexa proporciona más opciones a la hora de desarrollar la skill y de las diferentes formas en que se puede comunicar, proporcionando más expresiones, tonos de voz, etc.

Por otro lado, también se podrían utilizar otras técnicas de PLN aparte del método basado en lexicones utilizado, para dar un mayor porcentaje de acierto a la clasificación de las emociones y por tanto hacer que la respuesta que se devuelve al usuario sea más coherente.

Bibliografía

- [1] Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3), 436-465.
- [2] Rangel, I. D., Sidorov, G., & Guerra, S. S. (2014). Creación y evaluación de un diccionario marcado con emociones y ponderado para el español. *Onomazein*, (29), 31-46.
- [3] Chatterjee, A., Narahari, K. N., Joshi, M., & Agrawal, P. (2019, June). SemEval-2019 task 3: EmoContext contextual emotion detection in text. In *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 39-48).
- [4] Gupta, U., Chatterjee, A., Srikanth, R., & Agrawal, P. (2017). A sentiment-and-semantics-based approach for emotion detection in textual conversations. *arXiv preprint arXiv:1707.06996*.
- [5] Poria, S., Gelbukh, A., Hussain, A., Howard, N., Das, D., & Bandyopadhyay, S. (2013). Enhanced SenticNet with affective labels for concept-based opinion mining. *IEEE Intelligent Systems*, 28(2), 31-38.
- [6] Staiano, J., & Guerini, M. (2014). Depechemood: a lexicon for emotion analysis from crowd-annotated news. *arXiv preprint arXiv:1405.1605*.
- [7] Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4), 1093-1113.
- [8] Canales, L., & Martínez-Barco, P. (2014, October). Emotion detection from text: A survey. In *Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC)* (pp. 37-43).
- [9] Bruna, O., Avetisyan, H., & Holub, J. (2016, November). Emotion models for textual emotion classification. In *Journal of Physics: Conference Series* (Vol. 772, No. 1, p. 012063). IOP Publishing.
- [10] A. Yadollahi, A. G. Shahraki y O. R. Zaiane, «Current State of Text Sentiment Analysis from Opinion to Emotion Mining,» *ACM Computing Surveys*, vol. 50, no 2, pp. 1-33, 2017
- [11] Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, 98(45-60), 16.
- [12] Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In *Theories of emotion* (pp. 3-33). Academic press.
- [13] Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6), 1161.
- [14] Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14(4), 261-292
- [15] Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, 29(1), 24-54.

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

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

[18] Vásquez, A. C., Quispe, J. P., & Huayna, A. M. (2009). Procesamiento de lenguaje natural. *Revista de investigación de Sistemas e Informática*, 6(2), 45-54.

[19] Alswaidan, N., & Menai, M. E. B. (2020). A survey of state-of-the-art approaches for emotion recognition in text. *Knowledge and Information Systems*, 1-51.

[20] F.M. Plaza del Arco, C. Strapparava, L.A. Ureña López y M. Martín, «EmoEvent: A Multilingual Emotion Corpus based on different Events», *European Language Resources Association*, 1492-1498.

Anexo A. Manual de instalación

En este primer anexo se presenta el manual de instalación que hay que seguir en caso de querer utilizar el sistema creado.

La skill creada, como ya se ha explicado anteriormente funciona para Alexa de Amazon. Este AVI no se puede usar únicamente en un dispositivo físico (echo) de la compañía, si no que también existe la posibilidad de utilizarlo en cualquier dispositivo iOS (11.0+) o Android (6.0+)⁴⁰. Basta con entrar en la tienda oficial del sistema operativo e instalar la aplicación “Amazon Alexa” (Ilustración A.1.).

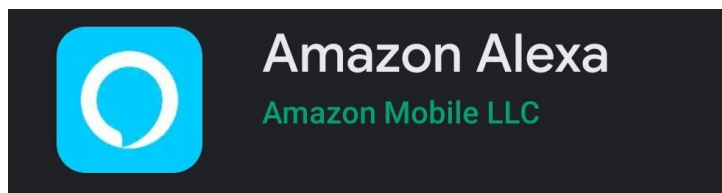


Ilustración A.1. Amazon Alexa App en Android

Para poder utilizar esta aplicación y la skill es necesario crear una cuenta en Amazon Alexa. Existen diferentes formas de habilitar la skill en nuestro dispositivo y poder utilizarla:

- Activar directamente mediante un comando de voz: “Alexa, activa yo te escucho”
- Desde el navegador, visitando la web de Amazon⁴¹, ingresando con nuestros datos. Posteriormente buscamos la skill ‘yo te escucho’ y pulsamos sobre ‘Activar’.
- Desde la aplicación de Alexa. Una vez entramos nos aparece algo similar a la Ilustración A.2., donde pulsaremos en el menú superior izquierdo. Posteriormente pulsaremos sobre ‘Skills y juegos’ (Ilustración A.3.). Por último buscaremos la skill ‘yo te escucho’ y pulsaremos sobre ‘Permitir su uso’ (Ilustración A.4.).

⁴⁰ <https://www.amazon.es/gp/help/customer/display.html?nodeId=GDHB98S32QPQ5H4Y>

⁴¹ <https://www.amazon.es/skills>

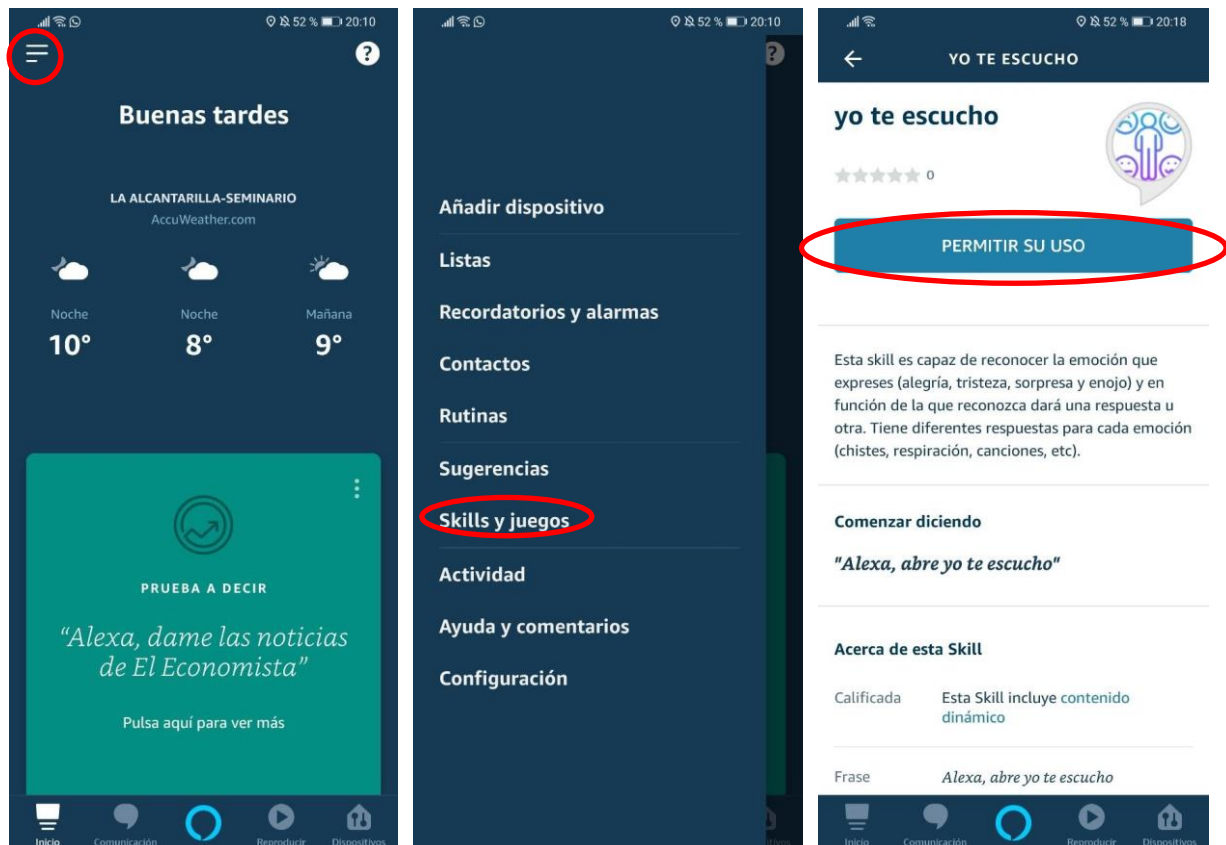


Ilustración A.2, A.3 y A.4. Proceso activar skill en app

Una vez hayamos seguido alguna de las opciones anteriores ya tendremos disponible para su uso la skill creada.

Anexo B. Manual de usuario

En este segundo anexo se presenta un manual del uso de la skill creada con la intención de resolver cualquier duda que pueda tener el usuario a la hora de utilizarla.

Para poder utilizar la skill es necesario que el usuario tenga un dispositivo 'echo' de Amazon Alexa o bien un dispositivo con sistema operativo iOS o Android.

El primer paso una vez se tiene activa la skill, proceso explicado en el Anexo A, consiste en abrirla. Para ello bastará con invocarla diciendo "Alexa, abre <nombre de la skill", en nuestro caso particular "Alexa, abre yo te escucho".

Si Alexa comprende correctamente lo que le hemos dicho, abrirá la skill, dándonos un mensaje de bienvenida y haciendo una pregunta similar a "¿cómo estás?". A esto contestaremos con lo que queramos, podemos decirle cualquier cosa que nos haya pasado, cómo nos sentimos en ese momento o incluso contestar con un simple 'sí' o 'no' si la pregunta lo permite. Esto será lo que la skill analizará para intentar reconocer una emoción.

Una vez hayamos dicho algo, existen dos opciones:

- La skill no ha escuchado nada. En este caso volverá a realizar la pregunta inicial.
- La skill ha escuchado lo que hemos dicho. En este caso, el texto se analizará en el clasificador de emociones que tiene integrado la skill y se dará una respuesta.

Una vez obtengamos una respuesta por parte de la skill existen dos opciones:

- Haya reconocido una emoción. En este caso nos dará una respuesta y nos preguntará por si queremos contarle algo más.
- No haya reconocido ninguna emoción, nos lleva al estado neutro. En este caso se volverá a preguntar algo similar a "¿algo más?", para que le digamos algo más que poder analizar.

Una vez llegados a este punto, podemos escoger dos caminos:

- Elegir finalizar la ejecución de la skill. Para ello basta con decir 'no' a su pregunta de si queremos contarle algo más.
- Decir algo más y volver a realizar el proceso explicado.

A continuación se muestran varios ejemplos de uso de la skill. El primer ejemplo, mostrado en la Ilustración B.1., se trata de una ejecución de las skill en la que se detecta tristeza por dos veces consecutivas.

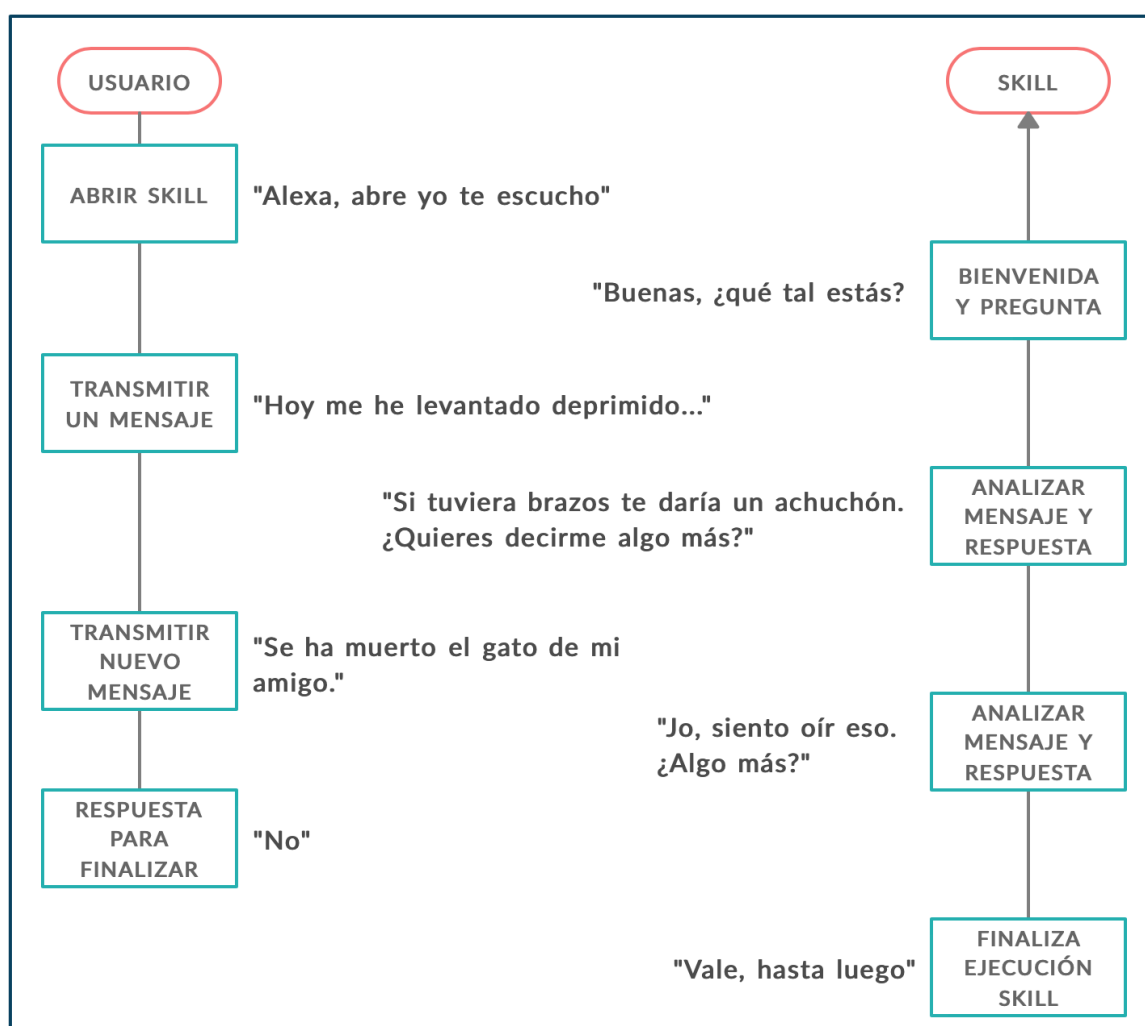


Ilustración B.1. Ejemplo flujo ejecución resultado tristeza

El segundo ejemplo, mostrado en la Ilustración B.2., se trata de una ejecución de la skill en la que se detecta alegría y posteriormente se cierra la skill contestando que no queremos contar nada más.

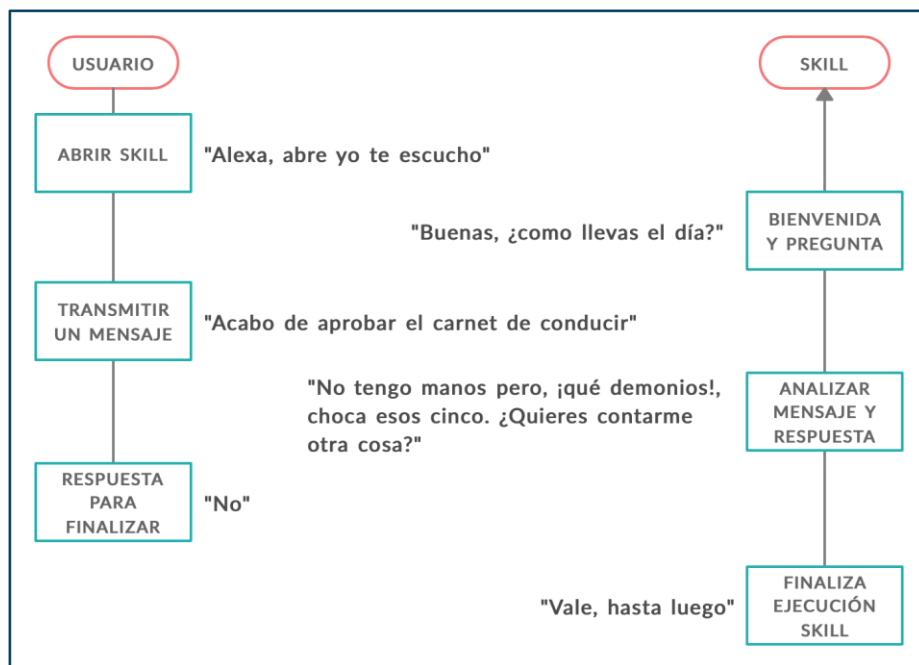


Ilustración B.2. Ejemplo flujo ejecución resultado alegría

El tercer ejemplo, mostrado en la Ilustración B.3., se trata de una ejecución de la skill en la que se detecta enojo.

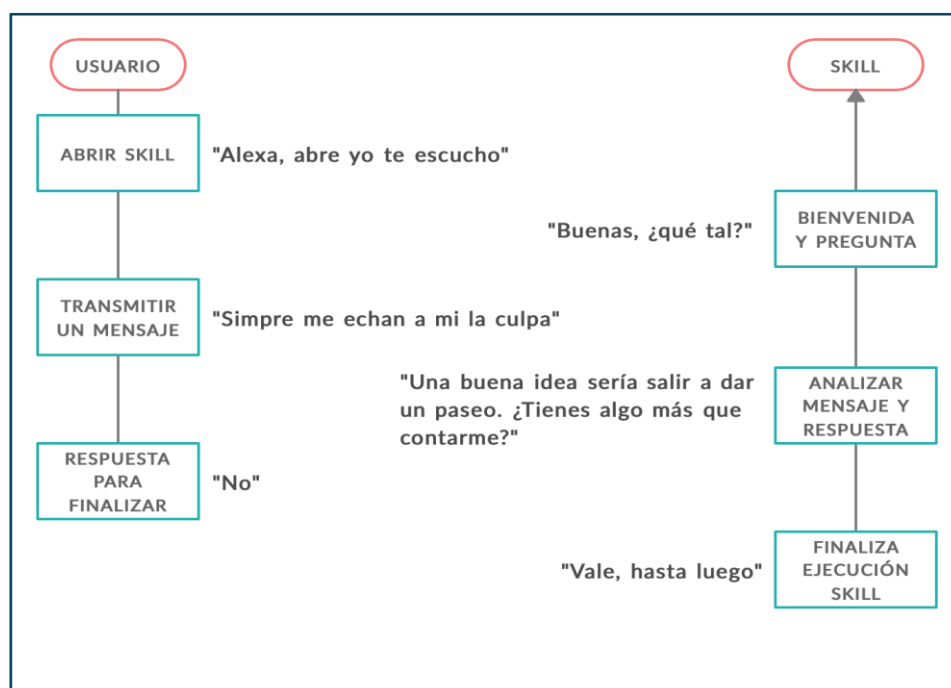


Ilustración B.3. Ejemplo flujo ejecución resultado enojo

El cuarto ejemplo, mostrado en la Ilustración B.4., se trata de una ejecución de las skill en la que se detecta sorpresa.

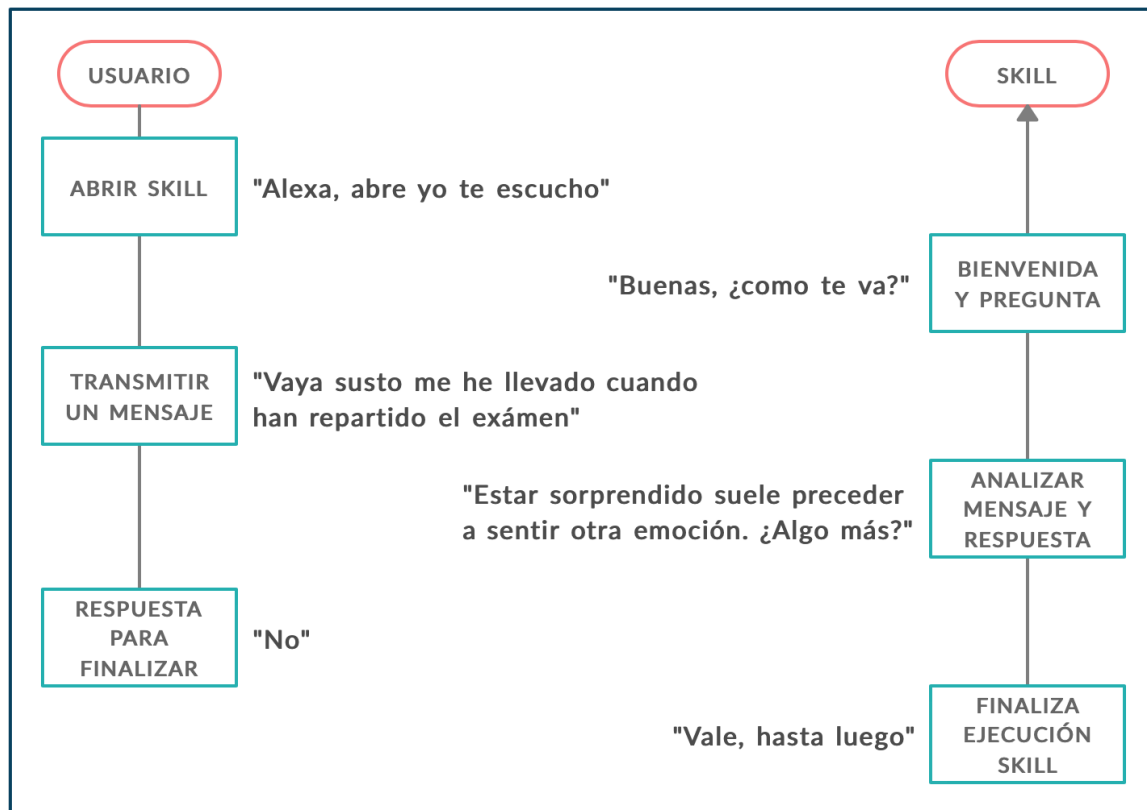


Ilustración B.4. Ejemplo flujo ejecución resultado sorpresa

El siguiente ejemplo, mostrado en la Ilustración B.5., se trata de una ejecución de las skill en la que se detecta el estado neutro y se vuelve a preguntar, obteniendo en la segunda ocasión tristeza.

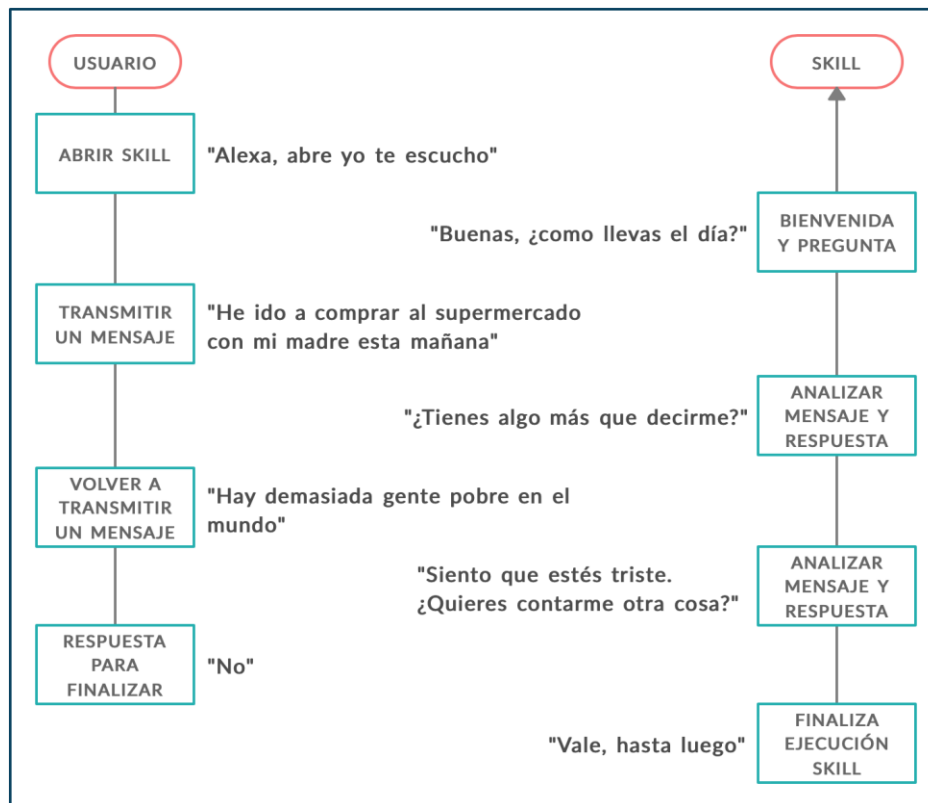


Ilustración B.5. Ejemplo flujo ejecución resultado neutro

En el último ejemplo, mostrado en la Ilustración B.6., la pregunta permite contestar con un simple 'sí' el cual se detecta como alegría.

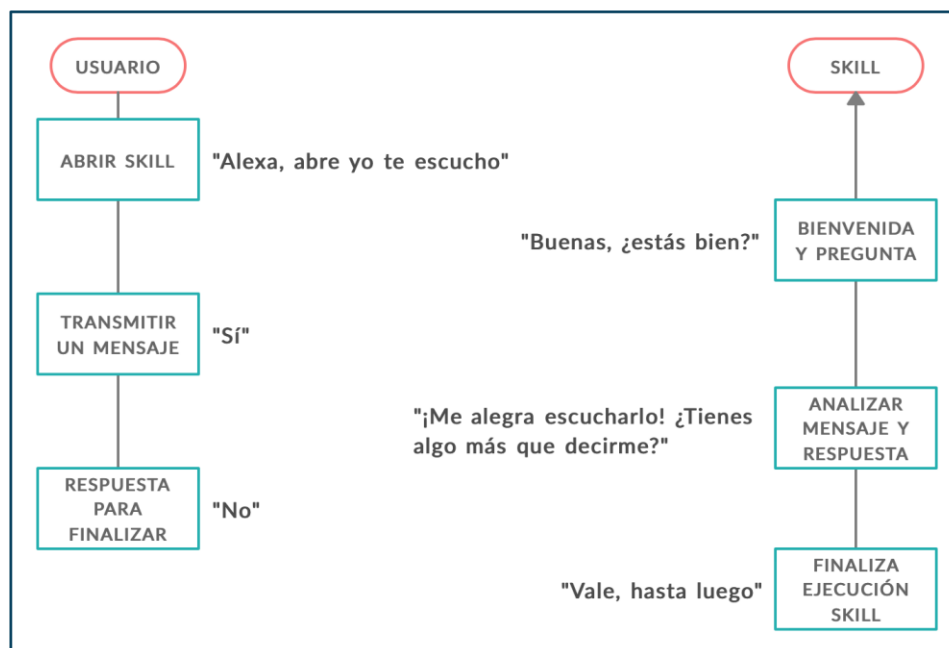


Ilustración B.6. Ejemplo flujo ejecución resultado alegría contestando 'sí'

Anexo C. Índice de ilustraciones

Ilustración 1.1. Miles de millones de AVI en uso por años

Ilustración 1.2. Millones de usuarios de Internet por lenguas

Ilustración 1.3. Diagrama de Gantt

Ilustración 2.1. Formas de utilizar Alexa

Ilustración 2.2. Partes de una skill

Ilustración 2.3. Orden para ejecutar una skill de Alexa

Ilustración 2.4. Intent y utterances

Ilustración 2.5. Funcionamiento de Alexa

Ilustración 2.6. Asistentes virtuales más utilizados en España

Ilustración 2.7. Dispositivos con AVI más utilizados

Ilustración 2.8. Dispositivos desde donde usuarios utilizan un AVI

Ilustración 3.1. Tareas dentro del Análisis de sentimientos

Ilustración 3.2. Emociones de Ekman

Ilustración 3.3. Emociones de Plutchik

Ilustración 3.4. Emociones de Russell

Ilustración 4.1. Ciclo de vida de una metodología tradicional

Ilustración 4.2. Ciclo de vida de una metodología ágil

Ilustración 4.3. Métodos ágiles

Ilustración 4.4. Ejemplo muro Kanban

Ilustración 4.5. Muro de kanban utilizado

Ilustración 4.6. Diagrama del sistema

Ilustración 4.7. Formato final LIWC

Ilustración 4.8. Formato final LIWC

Ilustración 4.9. Formato inicial Emolex

Ilustración 4.10. Formato final Emolex

Ilustración 4.11. Formato inicial SEL

Ilustración 4.12. Formato final SEL

Ilustración 4.13. Ejemplo ejecución resultado estado neutro

Ilustración 4.14. Ejemplo ejecución resultado emoción alegría

Ilustración 4.15. Ejemplo real de interacción del sistema final creado

Ilustración 4.16. Lenguaje de programación Python

Ilustración 5.1. Encuesta usuario 1.

Ilustración 5.2. Encuesta usuario 2.

Ilustración 5.3. Encuesta usuario 3.

Ilustración A.1. Amazon Alexa App en Android

Ilustración A.2. Proceso activar skill en app (I)

Ilustración A.3. Proceso activar skill en app (II)

Ilustración A.4. Proceso activar skill en app (III)

Ilustración B.1. Ejemplo flujo ejecución resultado alegría

Ilustración B.2. Ejemplo flujo ejecución resultado tristeza

Ilustración B.3. Ejemplo flujo ejecución resultado enojo

Ilustración B.4. Ejemplo flujo ejecución resultado sorpresa

Ilustración B.5. Ejemplo flujo ejecución resultado neutro

Ilustración B.6. Ejemplo flujo ejecución resultado alegría contestando 'si'

Anexo D. Índice de tablas

Tabla 1.1. Estimación temporal de las tareas del proyecto.

Tabla 1.2. Costes de personal

Tabla 1.3. Resumen de gastos totales del proyecto.

Tabla 2.1. Comparativa funcionamiento AVIs

Tabla 4.1. Plantilla historia de usuario

Tabla 4.2. Historia de usuario: “Procesado de lexicones”

Tabla 4.3. Historia de usuario: “Implementación front-end”

Tabla 4.4. Historia de usuario: “Implementación de clasificador de emociones (back-end)”

Tabla 4.5. Historia de usuario: “Capacidad para dar diferentes respuestas”

Tabla 4.6. Historia de usuario: “Evaluación del sistema”

Tabla 4.7. Batería de respuestas

Tabla 4.7. Batería de respuestas

Tabla 5.1. Usuarios que realizaron la encuesta de usabilidad