UNIVERSIDAD INTERNACIONAL SEK

FACULTAD DE ARQUITECTURA E INGENIERÍAS

Trabajo de fin de carrera titulado:

Desarrollo de una aplicación para el control parental de WhatsApp en dispositivos móviles Android

Realizado por:
Eduardo Francisco Caizaluisa Moreno

Director del proyecto:
Diego Fernando Riofrío Luzcando

Como requisito para la obtención del título de:
MÁSTER EN TECNOLOGÍAS DE LA INFORMACIÓN CON MENSIÓN EN SEGURIDAD Y REDES

DECLARACIÓN JURAMENTADA

Yo, EDUARDO FRANCISCO CAIZALUISA MORENO, con cédula de identidad número 171606165-8 declaro bajo juramento que el trabajo aquí desarrollado es de mi autoría, que no ha sido previamente presentado para ningún grado a calificación profesional; y, que ha consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración, cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la UNIVERSIDAD INTERNACIONAL SEK, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normativa institucional vigente.

Eduardo Francisco Caizaluisa Moreno
C.C.: 171606165-8

DECLARATORÍA

El presente trabajo de investigación titulado:

Desarrollo de una aplicación para el control parental de WhatsApp en dispositivos móviles Android

Realizado por:

Eduardo Francisco Caizaluisa Moreno

Como Requisito para la Obtención del Título de:

MÁSTER EN TECNOLOGÍAS DE LA INFORMACIÓN CON MENSIÓN EN SEGURIDAD Y REDES

Ha sido dirigido por el profesor

Diego Riofrío

Quien considera que constituye un trabajo original de su autor

Diego Riofrío
Director

LOS PROFESORES INFORMANTES

Frankie Catota

Daniel Riofrío

Después de revisar el trabajo presentado, lo han calificado como apto para su defensa oral ante el tribunal examinador.

Frankie Catota Daniel Riofrío

Quito, Junio 2018

DEDICATORIA

El presente trabajo es un tributo a los educadores que con su esfuerzo y paciencia son una guía para las nuevas generaciones que se forman. Los maestros realizan un trabajo enorme a pesar de las adversidades por la simple vocación de ser docentes, labor que algunas veces no es reconocido como tal por la sociedad.

Deseo hacer una mención especial a Diego Fernando Riofrío Luzcando que con sus conocimientos, apoyo y paciencia hizo posible la culminación de este proyecto que nació con el fin de ser un aporte para la sociedad en especial de nuestros niños.

AGRADECIMIENTOS

A mi esposa y a mis hijas por toda esa paciencia y apoyo que recibí durante todo este tiempo. Ellas compartieron conmigo ese deseo de superación y sed de conocimiento. Hoy les puedo decir que todo lo hice por ustedes.

A mis queridos maestros que con sus conocimientos inspiraron el presente trabajo, en especial al Dr. Diego Riofrío Luzcando quien con su guía, buenos consejos y sobre todo profesionalismo, ayudó a la conclusión exitosa de este proyecto.

RESUMEN

No hay duda que WhatsApp es una herramienta de comunicación muy utilizada entre los jóvenes en estos tiempos, sin mencionar que los teléfonos inteligentes son más comunes y accesibles entre ellos. WhatsApp es una aplicación que permite contactar con familiares, amigos, compañeros de estudio, entre otras personas. Cuando hacemos uso de esta tecnología estamos intercambiando información, en muchos casos sensibles, tal como datos personales, emociones, deseos. Sin embargo, cuando los mensajes que se intercambian tienen un grado de peligrosidad para los menores de edad, se vuelve importante monitorear esos mensajes.

Por otra parte, el ser humano ha estado fascinado por dotar de Inteligencia Artificial a los sistemas de cómputo para intentar emular el cerebro humano. Gracias a John Henry Holland padre de los algoritmos genéticos y pionero en este campo, hoy existen muchas aplicaciones que toman ventaja de investigaciones de este tipo, un ejemplo de ello son las aplicaciones que usan análisis lingüístico o gramatical para analizar lenguaje humano.

El presente trabajo plantea el desarrollo de una aplicación de control parental para teléfonos móviles *Android*, la cual toma de base los mensajes de emisor y receptor de WhatsApp, para posteriormente, mediante un proceso interno en el móvil, enviarlos a un servidor central donde son analizados y clasificados conforme al texto de la conversación que el menor está manteniendo y según las amenazas que se detecte es decir SEXO, DROGAS, *BULLYING*. Finalmente el aplicativo enviará una alerta informativa hacia sus padres para el respectivo control parental. Este proceso de clasificación se realiza con la ayuda de un algoritmo de análisis lingüístico.

ABSTRACT

There is no doubt that WhatsApp is a communication tool widely used among young people in these days, without mentioning that smartphones are more common and accessible among them. WhatsApp is an application that allows us to contact family, friends, and classmates, among others. When we use this technology we are exchanging information in many cases sensitive, such as personal data, emotions, and desires. When the exchanged message has a degree of danger for children, it is important to monitor them.

On the other hand, humans have been fascinated by equipping computer systems with artificial intelligence to try to emulate the human brain. Thanks to John Henry Holland, father of genetic algorithms, and a pioneer in this field, today there are many applications that take advantage of investigations of this type. An example of this, are applications that use linguistic or grammatical analysis to analyze human language.

The present work proposes the development of a parental control application for Android mobile phones, which takes messages from WhatsApp, then sends them to a central server, where they are analyzed and classified according to the text of the conversation that the child is maintaining, according to the threats that is detected SEX, DRUGS, BULLYING. Finally, the application sends an informative alert to their parents for the respective parental control. This classification process is carried out with the help of a linguistic analysis algorithm.

Índice General

Índice de figuras

Índice de cuadros

Formulas

CAPÍTULO I

INTRODUCCIÓN

El presente trabajo de investigación surge de la necesidad de ser un apoyo al control parental ante situaciones que pueden vulnerar la integridad a los menores de edad en el uso habitual de la tecnología. En este estudio a manera de introducción, se tocan temas relacionados con los delitos informáticos y la forma como estos afecta a los menores, muchas veces con la complicidad de las redes sociales. En ese sentido es trascendente analizar sobre la importancia del control parental y como este control ejercido hacia los hijos, puede ser apoyado con la ayuda de herramientas de software, ya que muchas actividades que realizan los menores de edad están relacionadas con la tecnología, por ejemplo, el uso de redes sociales, navegación por Internet, intercambio de mensajería, etc.

Actualmente, los dispositivos móviles son más accesibles y están fácilmente al alcance de la población de todas las edades. Estos dispositivos en su mayoría, utilizan *Android* como sistema operativo ya que está presente en el 86.1% de los dispositivos a nivel mundial según *Gartner* (2017). En este sentido es necesario realizar un análisis y una taxonomía a los aplicativos de control parental para dispositivos *Android* y sus funcionalidades, tanto a los existentes en el *Google Play Store* como los propuestos por la comunidad científica.

En el presente trabajo se plantea la creación de un aplicativo de Control Parental para el análisis de los mensajes y el envío proactivo de mensajes de alerta temprana como una solución a la problemática de los riesgos contra los menores de edad cometidos a través de *WhatsApp*, con el fin de informar a los padres sobre la presencia de un posible riesgo contra el menor. Para esto, nos apoyaremos del análisis de procesamiento de lenguaje natural mediante un algoritmo de clasificación semántico, con el cual se analizan los mensajes y el consecuente pronóstico sobre un posible riesgo dentro de las categorías definidas como SEXO, DROGAS y *BULLYING*.

1.1 El problema de investigación

1.1.1 Planteamiento del problema

En estos días la utilización de *WhatsApp* para comunicarse entre amigos y familiares es común en el diario vivir de las personas, sin embargo, se han presentado casos donde los menores de edad toleran de amenazas tales como acoso infantil, pornografía infantil, venta de sustancias estupefacientes y secuestros.

Según el portal web de la Policía Nacional del Ecuador[1] informa que se han reportado casos sobre las amenazas antes mencionadas. Cuenta la versión de un agente de la *UNASE* quien participó en un operativo, en el cual el extorsionador solicitaba cierta cantidad de dinero a una mujer a cambio de no atentar contra su vida ni las de sus familiares. En opinión del agente, según el *modus operandi* de los maleantes, estos buscan a sus víctimas a través de las redes sociales, en donde sacan la mayor información posible.

En el Ecuador, los delitos sobre medios electrónicos están tipificados por el Código Orgánico Integral Penal (2014) en los artículos Art. 169, 170, 171, 172 y en especial en el Art. 173 y 174 que tratan sobre delitos por medios electrónicos.

Un delito presente en estos días contra menores por medio de *WhatsApp* es el *sexting*, este, en el cual López, et. al. (2017) lo definen como "la actividad de enviar o recibir mensajes de texto, imágenes o videos con contenido sexual o erótico a través de teléfonos inteligentes u otros medios electrónicos". En este estudio, realizado en la Universidad Técnica de Manabí, de una muestra de 149 encuestados, se desprende que el 52.83% de los participantes de una encuesta sobre envío de contenido sexual, indican haber enviado mensajes sugestivos o de carácter sexual a otra persona sea pareja o amigos.

Otro grave problema que están sufriendo los menores durante la etapa escolar es el *bullying* o acoso escolar, explicado por Continente, et. al. (2010), como un abuso de poder repetitivo aplicado de forma violenta hacia un menor de edad. Según un informe de la *UNICEF* (2017), nueve de cada diez jóvenes están conscientes que el acoso escolar es un problema generalizado en sus comunidades y centros de estudio, y otros dos tercios indican haberlo experimentado personalmente.

Según la Encuesta sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los Hogares del Instituto Nacional de Estadística de España (2013), indica que el uso de los dispositivos móviles es cada vez más común entre la población infantil, así mismo se informa que el 50,9% de los niños españoles de hasta 11 años ya dispone un dispositivo móvil con tendencia creciente, y una cifra que crece hasta el 93,9% entre los adolescentes de 15 años de edad.

1.1.2 Diagnóstico

De acuerdo al estudio de García (2016), con respecto a las solicitudes de amistad que los menores aceptan en las redes sociales, el 71,3% de los menores ha indicado que se aseguran de conocerle y lo aceptan o lo eliminan. Dicha investigación indica que un 6.5% de la población investigada confiesa tener contactos de desconocidos en su libreta de contactos

[1] http://www.policiaecuador.gob.ec/sujeto-que-extorsionaba-mediante-mensajes-de-whatsapp-fue-detenido-por-la-unase/

del dispositivo móvil o redes sociales convirtiéndose en la oportunidad para que un delincuente entable algún tipo de relación con los menores.

Por otro lado, en el mismo estudio de García (2016) ante la pregunta de ¿tus padres conocen los usos que le das a Internet o a las redes sociales? el 19.7% de los jóvenes respondieron "Sí, lo revisan", un 60.1% "Sí, pero no lo revisan", el 16.7% "Sólo lo que me interesa", y "No saben nada" el 3.4%.

La columnista Galina Mursalieva del diario ruso Novaya Gazeta, publicó el 15 de mayo del 2016 una historia trágica en un extenso reportaje titulado "Grupos de la muerte", en donde informaba del suicidio de 130 niños desde noviembre de 2015 a abril de 2016, todos ellos pertenecientes a un grupo en redes sociales. La columnista publicó esta noticia para alertar a los padres a tomar precauciones hacia sus hijos antes de que sea demasiado tarde. Esta entrega de Galina Mursalieva, se convierte en el primer caso documentado en el mundo para denunciar el juego de la ballena azul, en donde se narraba las actividades que ocurrían en grupos de salas de chat como *VKontakte* o *HURRY ME AT 4:20* para alentar a los adolescentes a infringirse heridas o incluso a suicidarse. Dibujos de una silueta de ballena, videos de ballenas varadas en las orillas de las playas muriendo, canciones depresivas y de incitación al suicidio en el cual exhortaban a los jóvenes a tomar la fatal decisión.

En una noticia del periódico argentino La Nación (2017), se informó sobre el primer caso de suicidio de un adolecente de 14 años informado a través de redes sociales, específicamente este caso sucedió en la provincia de San Juan, a pesar que la victima había enviado mensajes solicitando ayuda los días previos al suicidio, sin embargo, nadie los tomo en cuenta. En cuanto a sus padres desconocían de lo que el joven estaba atravesando.

La tendencia hacia los delitos informáticos están en alza, así lo revela una investigación de la Oficina de Regulación de las Comunicaciones del Reino Unido (2015), en el año 2014 el 20% de menores entre la edad de 12 a 15 años utilizaba *WhatsApp*, frente al 23% en el 2015. Según la *Office for National Statistics* (2017), en el año 2016 hubo 458 delitos informáticos reportados contra adultos y niños frente al 884 en el 2017. Lo que demanda el estudio y desarrollo de tecnologías que mitiguen el problema.

1.1.3 Pronóstico

El desconocimiento que los padres tienen de la información que los menores manejan en *WhatsApp*, aporta a la materialización de las amenazas contra ellos sobre todo si no se cuentan con herramientas para comunicar a los padres sobre la información que manejan sus hijos en los dispositivos móviles, ya que la utilización de *WhatsApp* se vuelve más popular cada año.

1.1.4 Control de Pronóstico

Si los dispositivos móviles contaran con herramientas de control parental mejoraría el conocimiento de los padres con respecto al intercambio de información que hacen sus hijos en los dispositivos móviles. Estas herramientas deben ser capaces de emitir algún tipo de alerta cuando se detecte un delito.

1.1.5 Formulación del problema

El desconocimiento que tienen los padres frente a la información que manejan los menores en los dispositivos móviles aporta a la probabilidad de materialización de los riesgos contra ellos.

1.2 Objetivo General

Desarrollar una aplicación para dispositivos móviles *Android*, la cual ayude al control parental para la detección de riesgos contra menores de edad en *WhatsApp* a través del envío de alertas tempranas, utilizando herramientas de procesamiento de lenguaje natural.

1.3 Objetivos Específicos

- Definir una arquitectura cliente/servidor para el soporte todos los componentes diseñados para el aplicativo.
- Describir la estrategia de obtención de privilegios en el dispositivo móvil y la forma para acceder y obtener la base de datos de *WhatsApp* con el fin de obtener el insumo de datos para el aplicativo.
- Implementar un servicio que, por medio de escuchar en *Sockets*, reciba y procese la información que envían los diferentes dispositivos que tengan el aplicativo instalado.
- Diseñar y desarrollar un aplicativo móvil en *Android* que realice la recolección de los mensajes y los envíe al servicio.
- Implementar aprendizaje de máquina supervisada para clasificación de contenido, basado en un algoritmo conocido con la ayuda de la herramienta *Scikit-learn*.
- Implementar un mecanismo para el envío de alertas a los padres de familia suscriptos al servicio.
- Validar el modelo planteado mediante una simulación de conversaciones.

1.4 Justificaciones

En esta década donde la masificación y el avance de las tecnologías están en auge, gracias al fácil acceso a dispositivos móviles y su bajo costo, así como la disponibilidad de acceso

a redes de datos, desarrollamos nuestras actividades en tecnologías como estas y en las comunicaciones que ellas ofrecen.

En estos tiempos es fácil encontrar a más jóvenes usando y abusando de estas tecnologías. Sin embargo, no todo es bueno, los delincuentes han volcado sus esfuerzos para encontrar personas vulnerables que utilicen las redes sociales para contactarlos. Generalmente utilizan el engaño mediante la creación de perfiles falsos o la suplantación de identidad para llegar a las víctimas y así cometer delitos.

Los medios electrónicos de noticias informan con regularidad de casos de extorciones, como es el caso de una publicación de Jiménez (2017), del medio electrónico *Excelsior* de México, informó sobre un agresor que extorsionó a la familia de una víctima, a quien previamente le habían robado su celular. Los extorsionadores se comunican enviándoles mensajes, fotos, audios y videos para dar fe de vida de la víctima y de su secuestro, de esta manera intentaron obtener grandes cantidades de dinero por la liberación de la víctima.

Gracias a que la información es globalizada, la difusión de modas, costumbres y tendencias también lo es. La facilidad de intercambio de información y comunicación por redes sociales por parte de los menores sin ningún tipo de filtro o restricción, se puede convertir en un serio problema si ellos son asechados por delincuentes, lo que puede desencadenar en el cometimiento de delitos hacia menores si existe poco o nulo control parental por falta de mecanismos o por desconocimiento.

Los menores también son víctimas de amenazas como el *Sexting*, pornografía infantil, acoso escolar o comúnmente llamado *Bullying*, los mismos que pueden ser cometidos por medio de *WhatsApp*, estas amenazas pueden causar daños psicológicos e incluso físicos en el menor, si no se detecta y detiene a tiempo.

Bajo este contexto, se vuelve urgente la necesidad que los aplicativos cuenten con mecanismos para alertar sobre la presencia de amenazas, que, si bien existen aplicaciones para el control parental, estas no llegan al nivel de analizar y clasificar el contenido marcándolo como nocivo de ser el caso y posteriormente alertar de manera temprana.

En este sentido, la propuesta para este proyecto de tesis es desarrollar un aplicativo que alerte de forma temprana a los padres, después de realizar un análisis y clasificación del contenido de los mensajes, marcándolos como nocivos en el intercambio de información de los dispositivos móviles a través de *WhatsApp*. Estas alertas pueden ser un mecanismo de gran ayuda para el control parental reduciendo proactivamente los riesgos contra los menores.

1.4.1 Teórica

Desde el punto de vista de la academia, la investigación serviría tanto a nivel técnico como científico para investigaciones de inteligencia artificial, procesamiento de lenguaje natural, clasificación y pronósticos de datos.

En el contexto del ámbito científico, este proyecto recoge varios ámbitos del conocimiento que se citan a continuación:

- Conocimiento en el ámbito académico científico que promoverán debates sobre las técnicas de análisis de texto, clasificación y explotación de información implementadas con la finalidad de aplicarlas en otros campos.
- Metodologías, arquitectura de *software* adoptadas para lograr plasmar el aplicativo y volverlo funcional acorde a las necesidades de los usuarios.
- En la minería de datos con algoritmos de análisis de textos que identifiquen tendencias y patrones de comportamiento de la información con la finalidad de tomar decisiones.
- Análisis de patrones de textos para determinar sentimientos positivos o negativos y así prevenir situaciones que lleven a tragedias.

1.4.2 Metodológico

Desde el punto de vista metodológico se utilizó el método exploratorio para la indagación y análisis de contenido dentro del dispositivo móvil y de las bases de datos de *WhatsApp*. Desde el punto de vista técnico, mediante un proyecto de desarrollo, se creó una aplicación móvil, para lo cual se utilizaron técnicas de prueba y error durante el desarrollo, y de validación empírica para las pruebas y análisis de los resultados. Además, para el desarrollo de la aplicación de Control Parental, se investigó información de mecanismos de recolección y procesamiento de datos en un dispositivo con sistema operativo *Android*.

1.4.3 Práctico

En cuanto al punto de vista práctico, esta solución está diseñada para ayudar al control parental, mediante la emisión de alertas tempranas, esto quiere decir antes de que las amenazas contra los menores ocurran. Para que esto ocurra, el aplicativo debe recolectar datos de *WhatsApp* para procesarlos, clasificarlos y analizarlos dependiendo si ciertos datos corresponden a un posible peligro. Bajo este contexto, se debe tomar en cuenta que es importante la identificación y adquisición de la base de datos de *WhatsApp*, puesto que es el insumo básico para el análisis de información.

1.5 Marco Teórico

Existen hechos violentos que se realizan contra personas y/o bienes, para poner un nombre normalmente aceptado, los nombraremos como delitos. Así, "los delitos violentos son sucesos negativos, vividos de una manera brusca que generan terror o indefensión, ponen en peligro la integridad física y psicológica de una persona y dejan en la víctima tal estado emocional que es incapaz de afrontarla con sus recursos psicológicos habituales" (Kilpatrick, Saunders, Amick-McMullan, Best, Veronen y Jesnick, 1989).

Según el Instituto Nacional de Estadística y Censos (2014), en una noticia publicada en mayo de 2014, afirma, en base a sondeos, que 12 millones de ecuatorianos tienen un *Smartphone*, esto representa un alto porcentaje en el global de la población que según el Banco Mundial (2018) proyecta una población de habitantes de 16.39 millones para el año 2016 en el Ecuador, lo que hace suponer una proporcional creciente en el uso de los dispositivos móviles en el país.

El acceso a Internet mediante el uso de un dispositivo móvil o *Smartphone*, además de tener beneficios de comunicación también trae vulnerabilidades a las personas frente a la delincuencia. Hoy en día los delincuentes encausan sus esfuerzos para cometer delitos del tipo informático, como lo menciona Ojeda-Pérez (2010), producto de la masificación en el uso de la Internet también ha aumentado el riesgo del mal uso que le demos a este medio por el que los delincuentes tratarán de incursionar de una u otra manera para realizar fraudes, estafas y extorción.

La violencia en redes sociales, como por ejemplo, el conocido juego de la ballena azul, que fue mencionado por primera vez por Galina Mursalieva (2016) y analizado por D'Adderio (2017), encuentran que a las víctimas se les presenta una serie de desafíos que buscan atentar con su integridad; tales como, tatuarse, ahogarse, cortarse, lanzarse desde una altura considerable para finalmente pedir el suicidio del jugador.

En el presente estudio se abordan los temas de *bullying*, el ciberacoso, así como los delitos informáticos con la finalidad de dar una solución a esta problemática, aportando una mejora en el control parental para minimizar los riesgos, y evitar que las amenazas se materialicen.

1.5.1 Marco Conceptual

En este apartado se tratarán los delitos informáticos y el ciberacoso en menores. Se tratarán también los aplicativos de Control Parental como herramientas de ayuda a los padres y las funciones básicas que un aplicativo de este tipo debe tener.

Estos aplicativos analizados están basados en el Sistema Operativo *Android*, ya que están presentes mayoritariamente en el mercado de los dispositivos móviles. Es importante indicar que algunos de los aplicativos sobre control parental basados en *Android*, utilizan procesamiento de máquina para analizar los mensajes en función de determinar patrones inusuales para la detección de algún tipo de amenaza en los textos de los mensajes. Para profundizar el tema de aprendizaje de máquina, se revisarán conceptos sobre Procesamiento de Lenguaje Natural, Algoritmos Bayesianos del tipo *Naíve Bayes* y clasificación de textos. Todo esto apoyado de la librería *Scikit Learn* la cual abstrae la complejidad matemática de los algoritmos utilizados.

1.5.1.1 Delito Informático

Según Valdés (1987), en su libro sobre derecho informático, define al delito informático como una forma antijurídica en las que se tiene a las computadoras como principal instrumento o fin para cometer un delito. Como por ejemplo la suplantación de identidades en las redes sociales, la pornografía infantil, el *Bullying* informático, etc.

Valdés (1987) define algunas características para los delitos informáticos. Entre las más relevantes tenemos:

- Solamente un determinado número de personas con conocimientos técnicos pueden llegar a cometerlas.
- Provocan pérdidas económicas a los afectados y producen beneficios a aquellos que los realizan.
- No es necesario presencia física para que puedan llegar a consumarse.
- Son muchos los casos y pocas las denuncias.
- Son actividades muy sofisticadas y más frecuentes de lo que se piensa.
- Debido a su complejidad técnica, presentan enormes dificultades para su comprobación.
- Cada año tienden a proliferar. Cada vez requieren de una regulación y judicialización más eficiente.

1.5.1.2 Bullying o Acoso Escolar

Conforme a la definición realizada por Vallejo et. al. (2017), el *Bullying* o acoso se refiere a conductas reprochables que pueden llegar a generar drásticas consecuencias entre los afectados, generalmente padecido por los menores en etapas escolar o secundaría, donde no solo los menores son los involucrados sino el acosador, así como también los espectadores o testigos, lo que denominan los autores como el "triángulo del *Bullying*".

1.5.1.3 Ciberbullying

El *Ciberbullying* (ciberacoso), según Vallejo (2017), se asemeja al *Bullying*, pero en este caso se utiliza la tecnología como un medio para el acoso, un ejemplo de esto es la publicación de imágenes ofensivas o de carácter sexual referentes a la víctima, lo que se puede convertir en un problema mucho peor cuando estos mensajes o imágenes son difundidos a un número considerable dentro de la red social. En el ciberacoso, según el autor, deben concurrir los tres elementos básicos del acoso escolar: intencionalidad, repetición, y desequilibrio de poder; añadiendo la peculiaridad de que se produce a través de las *TIC*[2]*s*, como medio para recibir y ejercer el maltrato.

1.5.1.4 Menores de Edad

Biológicamente hablando la adolescencia, de acuerdo a Shutt-Aine, J., & Maddaleno, M. (2003) comienza en la mayoría de personas al final de los 11 o 12 años, todo comienza mediante un rasgo característico se da por el acelerado crecimiento que se inicia en la pubertad, cuando principalmente se dan cambios biológicos, y psico-emocionales, en esta etapa se da los crecimientos de forma más rápida donde se experimentan cambios importantes desde el punto de vista físico, como son el crecimiento, el aumento de peso y la aparición de características sexuales secundarias como bello en los órganos sexuales, crecimiento de senos o aparición de la menstruación en las mujeres.

Según el Código de la Niñez y Adolescencia (2002), un niño o niña se considera como tal antes de la adolescencia en la edad de 12 o 13 años y un adolecente se considerará como tal antes de cumplir los 18 años de edad, ante la ley, ellos son considerados como menores de edad.

1.5.1.5 Control Parental

Stattin y Kerr (2000) sugirieron que los padres pueden llegar a tener un alto conocimiento de lo que hacen sus hijos mediante un seguimiento efectivo de las actividades que realizan los menores, lo que estos autores llamaron en su artículo "conocimiento de los padres". Este conocimiento refleja la medida en que los padres activamente buscan información sobre sus hijos, esto es: paradero, actividades realizadas y compañeros frecuentes como amigos e incluso padres de amigos. Este "conocimiento de padres" es similar a lo que otros autores como Fletcher et. al. (2004) llaman como "monitoreo parental". Stattin y Kerr conceptualizaron el control parental como la medida en que los padres requieren que los

[2] Tecnologías de la Información y Comunicación

adolescentes obtengan su permiso antes de salir e insisten en que se les informe sobre el paradero, las actividades y las amistades que frecuentan sus hijos.

Fletcher et. al. (2004) llaman al "conocimiento de padres", como "monitoreo parental". Además, argumentan que los padres también pueden aumentar este conocimiento sobre las actividades de los adolescentes ejerciendo altos niveles de control sobre los niños. Los mismos autores aseguran que cuando los padres exhiben niveles más altos de control parental sobre las actividades de sus hijos y compañeros, se puede evitar comportamientos nocivos en los hijos, minimizando así, el consumo de sustancias estupefacientes y hábitos delincuenciales. Los malos hábitos deberían operar independientemente en la medida en que los padres tienen más conocimiento y ejercen el control parental de manera adecuada.

1.5.1.6 Aplicativo de Control Parental

Los aplicativos de control parental, según Marcelo, J. F. y Martín, E. (2010) son programas de software que han sido diseñados para controlar y monitorear dispositivos que generalmente están en manos de menores de edad y que los padres tienen la intención de supervisar el uso que los menores realizan en el equipo. En estos programas de software se definen reglas que rigen en un dispositivo para bloquear, monitorear o controlar la ejecución de tareas dentro del mismo, y que pueden ser perjudiciales para el menor, estos y otros temas son detallados en el capítulo 2 donde preliminarmente se encontró que los aplicativos de control parental se clasifican en dos grandes grupos: los aplicativos que realizan monitoreo y los aplicativos que restringen.

Los aplicativos de monitoreo son aquellos que se encargan de vigilar el uso que el menor hace en el dispositivo y los tipos de aplicativos que restringen. Es decir, se encargan de bloquear por ejemplo ejecución de aplicaciones, descargas, compras en el *Play Store*, etc.

Entre las principales características que los aplicativos de control parental ofrecen tenemos las siguientes.

- Rastreo y localización de lugares visitados.
- Actividades desarrolladas en el dispositivo.
- Monitorear sitios de navegación.
- Monitorear mensajería.
- Monitorear contenido multimedia.
- Monitorear *WI-FI, Bluetooth, modem.*
- Restringir instalación y ejecución aplicaciones.
- Ejecución, descargas, instalación de aplicaciones.
- Configuración del sistema y de aplicaciones.

- Navegación en Internet.
- Llamadas entrantes / salientes a números autorizados.

1.5.1.7 Funciones Básicas de un Aplicativo de Control Parental

Conforme al sitio web welivesecurity[3] de *ESET*, la cual es una empresa líder mundial en antivirus y protección de ordenadores, existen 5 medidas que se pueden tomar en cuenta para el control parental:

- Roles de usuario.
- Restricciones de acceso por categoría.
- Bloqueo de páginas *web.*
- Registro de actividad.
- Control de aplicaciones y de tiempo.

Según el mismo sitio web de ESET[4], resume las principales características que se requieren para garantizar la seguridad de los hijos en los dispositivos móviles:

- Control de aplicaciones.
- Localizador *GPS.*
- Control *Web.*
- Limita el tiempo de juego.
- Mensajes de papá y mamá.

1.5.1.8 Desarrollo de aplicativos móviles

La diferencia entre un aplicativo móvil y una aplicación de escritorio es la capacidad de los recursos debido a que en un celular son más limitados por su tamaño (memoria, batería). Por otro lado, está el tamaño de la pantalla, ya que sobre un área tan pequeña debe administrarse efectivamente los componentes y el diseño de las pantallas.

En el mercado de dispositivos móviles uno de los sistemas operativos se llama *iOS*, de la multinacional *Apple. Inc.*, el cual según *Garner* (2017) es el segundo operativo utilizado a nivel mundial detrás de *Android*. Este sistema operativo se deriva de *MacOS*, el mismo que está basado en sistemas *unix* cuya versión es *Darwin* BSD[5]. El desarrollo de aplicativos para *iOS* se lo realiza mediante el lenguaje *Objetive-C*, sin embargo, *Apple. Inc.* lanzó un nuevo lenguaje denominado *Swift* en el año 2014, el cual según García et. al. (2015), trae mejoras en sintaxis, reducción de líneas de código, mejoras en estructuras de datos. La

[3] https://www.welivesecurity.com/la-es/2015/09/25/supervisar-herramienta-de-control-parental/
[4] https://www.eset.es/productos/control-parental-android/
[5]https://developer.apple.com/library/content/documentation/Darwin/Conceptual/KernelProgramming/Architecture/Architecture.html

interfaz de desarrollo se denomina *xcode* la misma que plantea el uso de *Storyboards* (guiones gráficos) y eventos para las pantallas de un aplicativo en *iOS*.

De acuerdo a Google (2010), *Android* es un *software* para dispositivos móviles, basado en el sistema operativo Linux cuya arquitectura[6] está compuesta por el *Kernel* de *Linux*, *Hardware Abstraction Layer* (*HAL*), Librerías Nativas, *Android Runtime*, *Android Framework*, Aplicaciones el mismo que se trata de una máquina virtual que interpreta el código binario compilado, con el fin de optimizar al máximo los recursos físicos del equipo.

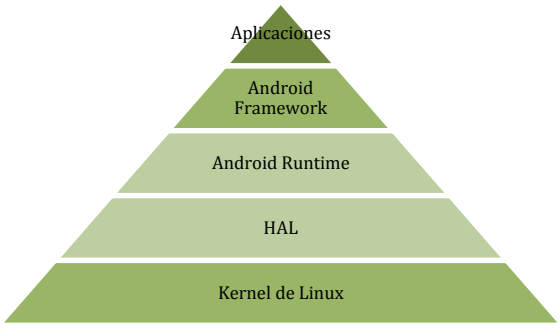

Figura 1. Arquitectura en capas de un sistema *Android* basado en source.android.com[6]

El lenguaje por defecto, en el que se desarrollan aplicaciones para *Android*, es *JAVA* debido a que su máquina virtual *DALVIK* originalmente fue desarrollada en este lenguaje. Cabe destacar que se podría desarrollar aplicaciones mucho más eficientes en *C++*, gracias a que *Android* permite compilación en código nativo con la librería *NDK*[7], un ejemplo de aquello son los videojuegos.

Para el desarrollo de una aplicación en *Android*, principalmente es necesario contar con el *SDK* (*Software Development Kit*), el cual contiene todas las librerías para que un proyecto de software en *Android* funcione. Por otro lado, se requiere de una herramienta *IDE* (*Integrated Development Environment*) para administrar, editar, depurar y compilar los *scripts* del proyecto, los *IDEs* más utilizados[8] son *Android Visual Studio*, *NetBeans* y Eclipse.

Para implementar un proyecto de *Android* y hacer pruebas de ejecución, se puede utilizar el simulador denominado *AVM* (*Android Virtual Machine*) o directamente con un dispositivo móvil conectado al computador de desarrollo mediante un cable *USB*.

[6] https://source.android.com/devices/architecture/
[7] https://developer.android.com/ndk/
[8] https://academiaandroid.com/ide-entornos-integrados-de-desarrollo-para-android/

1.5.1.9 Aprendizaje de máquina

Según Guyon, I. (2008) el proceso de aprendizaje de máquina consiste en tener una base de datos grande que se coloca como entrada a un algoritmo, el cual se ejecuta dentro en una máquina entrenada para obtener una salida, en el proceso de utilización el usuario realiza preguntas o consultas a la máquina y obtiene respuestas. Estos algoritmos permiten principalmente realizar: clasificaciones de problemas (texto), predicciones, regresiones y agrupamiento de datos.

Guyon (2008) también clasifica el aprendizaje de máquina en 4 métodos: modelos lineales, métodos de *Kernel*, redes neuronales y árboles de decisión, de estos métodos se forman 4 categorías de máquinas que pueden aprender a resolver un problema en particular.

- Métodos Lineales y de *Kernel*
  El objetivo de los modelos lineales, según Guyon (2008), es construir una función lineal tal que permita realizar predicciones cuando se da una nueva entrada. Con respecto al método *kernel*, éste consiste en transformar las entradas tomando en cuenta el peso de cada una de las filas en los datos de entrenamiento.

- Redes Neuronales
  Astray et. al. (2010), en su estudio sobre redes neuronales, se refieren a la interconexión de neuronas o denominadas nodos o grupo de nodos que forman una red simulando la interconexión de las neuronas de un cerebro humano. Estas neuronas cuentan con una entrada la cual produce una salida, y a su vez, esta salida puede convertirse en la entrada a de otra neurona formando así una red neuronal extensa.

- Árboles de Decisión *RF (Random Forest)*
  Los árboles de decisión, según Rokach y Maimon (2008), son una estructura de nodos con un nodo raíz origen. Esta estructura se compone de nodos, vectores, flechas y etiquetas, de esta forma se sigue un camino en base a diferentes factores para tomar una decisión.

- *Naíve Bayes*
  Abad-Grau et. al. (2007) y Mendoza et. al. (2011) lo consideran como un clasificador bayesiano, el cual utiliza una representación de documentos livianos llamada bolsa de palabras o *bag-of-words* y reúne una co-ocurrencia de términos basadas en la aproximación. Este clasificador es tratado más a detalle a continuación.

1.5.1.10 *Naíve Bayes*

El algoritmo *Naíve Bayes*, también conocido como aprendizaje bayesiano inspirado en la

biología y basado en la teoría Bayesiana según Escolano Ruiz, F. (2003). De acuerdo a Martínez et.al. (2009) Naíve Bayes, es un clasificador que predice eventos basado en sus características, el resultado producido se lo conoce como pronóstico. Para producir el pronóstico, el clasificador genera un árbol de decisión para pronosticar posibles resultados en base a la probabilidad de que una instancia dada pertenezca a un conjunto de clases, Oded y Rokach (2005).
La fórmula de Naíve Bayes estima la probabilidad a posteriori $P(A_i|B)$ resultante, por otro lado B es una clase cualquiera donde se conoce que $P(B|A_i)$ es la probabilidad de B en el conjunto de A_i o en otras palabras, representa la probabilidad de que ese evento ocurra en A, donde A es un conjunto de clases conocidas $\{A_1, A_2, A_3, \ldots, A_n\}$ y mutuamente excluyentes.

$$\boldsymbol{P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}}$$

Fórmula 1. Función Probabilidad Bayesiana

Hoy en día, existen 3 variaciones del Algoritmo Bayesiano, los cuales son: *Gaussian Naíve Bayes, Multinomial Naíve Bayes y Bernoulli Naíve Bayes*, según Mendoza et. al. (2011). Estas variaciones han permitido obtener mejores resultados adaptados a las circunstancias que se requieran.

1.5.1.11 Multinomial *Naíve Bayes*

De acuerdo a Hernández (2009), el modelo *MultinomialNB* es un tipo de clasificador Multinomial, basado en la aplicación de la regla de *Bayes* conocido por su naturaleza dicotómica (Dos valores Éxito, Fracaso), el Multinomial realiza evaluación de n sucesos y en cada evaluación determina el éxito en cada seceso evaluado. Los cuales sirven para una larga escala de problemas de clasificación e identificación de características basados en la probabilidad de ocurrencia de múltiples categorías.

Por otro lado Hernández explica que en un modelo Multinomial se genera un arreglo con una serie de características $p = (p_1, \ldots. p_n)$ que representa la frecuencia que un evento p_i puede ocurrir, por otro lado tenemos otro arreglo que representa el número de ocurrencias observadas en un problema a resolver $x_i = (x_i, \ldots, x_n)$.

Repasando como *MultinomialNB* computa la probabilidad de ocurrencia de una clase para un documento dado basado en la fórmula descrita por McCallum, A., & Nigam, K. (1998), se tiene un conjunto de clases denotado por *X1, X2, X3.*. y *x* como una clase en particular, *n* viene a ser el tamaño del arreglo, entonces *MultinomialNB* asigna un documento de prueba x_i, a la clase que tiene la más alta Probabilidad *P* según el número de ocurrencias en una matriz.

$$P(X_1 = x_1, \quad X_2 = x_2, X_3 = x_3 \ldots) = \frac{n!}{x_1! * x_2! * x_3! * \ldots} p^1 * p^2 * p^3 * \ldots$$

Fórmula 2. Multinomial Naíve Bayes descrita por McCallum, A., & Nigam, K. (1998)

Scikit-learn[9] implementa el algoritmo *Naíve Bayes* para datos distribuidos multinomiales, que es una de las variantes de los dos algoritmos clásicos *Naíve Bayes* usados para clasificación de texto, es decir el Algoritmo *Gausian Naíve Bayes* y el *Bertoulli Naíve Bayes*.

1.5.1.12 Aprendizaje de Lenguaje Natural

Según Moreiro y Rodríguez (1999), el lenguaje natural es aquel conjunto de signos y símbolos orales, escritos por medio de los cuales los seres humanos se comunican entre sí. Por otro lado, en el proceso de lenguaje natural de forma computacional para lograr que el análisis funcione, es necesario realizar una indexación automatizada de palabras adyacentes para identificar a un documento por un conjunto de palabras claves representativas de su contenido. En el ámbito del análisis del lenguaje natural se realiza mediante el análisis de:

- análisis morfológico,
- análisis sintáctico,
- análisis semántico,
- análisis fonológico,
- y. una combinación entre análisis morfológico y sintáctico llamado morfosintáctico.

Se puede resumir a la indización automatizada como el uso de computadoras para extraer o asignar términos indexados sin intervención humana, una vez que se han establecido programas o algoritmos relacionados al procedimiento.

De acuerdo a Rodríguez (2007) el reconocimiento de texto se da por lo que él autor denomina como "*Información Semántica*", la misma que se compone de: palabras clave, redes semánticas y ontológicas. Adicional a esto, Graña Gil, J. (2002) reconoce como una característica importante la realización de "*análisis sintáctico*", para el análisis de lenguaje natural además conjuntamente con la etiquetación.

Los documentos son representados por medio de palabras claves en un *bag-of-words* G. Salton (1983). Para esto se genera un vector por cada documento en donde cada posición del vector representa una palabra clave del documento. Para Arias Figueroa, M. A. (2016), quien denomina a *bag-of-words* como un grupo de descriptores, una manera de representar a un grupo de palabras clasificadas según su peso en función de la ocurrencia en un documento.

[9] http://scikit-learn.org/stable/modules/naive_bayes.html

Conforme a Escolano Ruiz, F. (2003), el análisis de lenguaje natural se clasifica en dos métodos, el morfológico que trata sobre las reglas en la composición de palabras de un lenguaje y el sintáctico que trata sobre gramáticas, árboles sintácticos, y su conexión mediante una relación sintáctica propia del lenguaje.

1.5.1.13 Python

Montoro (2013) en su libro, define a *Python* como un lenguaje de programación basado en C creado a los finales de los años 90 en el centro de investigaciones de ciencias matemáticas, *Centrum Wiskunde & Informática* de Ámsterdam, bajo el liderazgo del investigador holandés Guido van Rossum.

Python fue diseñado para ser un lenguaje de alto nivel, interpretado y multipropósito, el cual puede ser implementado en plataformas como *Linux, Windows y macOS*.

El lenguaje de programación *Python* es un lenguaje por excelencia en el campo de la academia, como lo cita Challenger-Pérez, I. et.al. (2014), *Python* ha sido seleccionado por el prestigioso Instituto de Tecnología de *Massachusets* (*MIT*) para impartir clases en ciencias de la computación. En el campo de la ciencia es ampliamente utilizado en el *CERN* (Organización Europea para la Investigación Nuclear), y ampliamente utilizado por científicos en especialidades como Física, Bioinformática, Astronomía, Neurofisiología, etc. ya que cuenta con librerías como *SciPy* y *NumPy* por sus siglas *Python* científico y *Python* numérico respectivamente.

1.5.1.14 *Scikit Learn*

Según Pedregosa et. al (2011), *Scikit Learn* es una librería basada en *Python* (bajo una licencia *BSD*) que permite la implementación de un amplio rango de algoritmos de aprendizaje supervisados y no supervisados, provechosamente *Scikit Learn* abstrae la complejidad matemática implícita que tiene la implementación de los algoritmos de aprendizaje en código de *software*. Esta librería posee un abanico de algoritmos que se pueden aplicar para análisis de datos, minería de datos e inteligencia artificial.

Casos reales de implementación los encontramos en aplicaciones como *Clasificación de actividades humanas en tiempo real a partir de representaciones de esqueleto de Aguado* de Corman, A. (2016), *Predicción de la toxicidad y de la actividad antimicrobiana a partir de la secuencia aminoacídica* de Mariño Solís, R. (2017), *Desarrollo de un sistema distribuido para la digitalización y procesamiento de cheques usando algoritmos de reconocimiento de dígitos manuscritos en la empresa Decisión CA* de Benalcázar, S. et. al. (2016), *Aplicación de ciencia de datos para la creación de software predictivo de*

morbimortalidad materna en México de Domínguez Domínguez, R. (2017).

1.6 Adopción de una Perspectiva Teórica

En primer lugar se plantea el desarrollo de un Aplicativo basado en *Android*, ya que este sistema operativo cuenta con el mayor mercado en teléfonos inteligentes a nivel mundial. Según *Gartner* (2016) está presente en el 86.2% del total de dispositivos inteligentes en el mercado. Por otro lado, *Android* cuenta con un conjunto de herramientas de libre descarga disponibles para el desarrollo de aplicaciones como son *IDE*, librerías, soporte, actualizaciones constantes, comunidad, foros, *blogs* y documentación.

Frente al problema de investigación planteado donde se observa la existencia de riesgos contra los menores y el desconocimiento que tiene de esas amenazas por parte de los padres. Por lo que el presente trabajo plantea un aplicativo para analizar los mensajes que se envían a través de *WhatsApp* con la finalidad de enviar alertas tempranas para mejorar el control parental en los menores. El análisis de los mensajes es importante ya que en ellos existe información crucial que debe ser explotada con la finalidad de identificar amenazas, de acuerdo a Lucio, P., & Vicente, J. (2018), en la escena del crimen un médico forense puede usar varios recursos del tipo tecnológico para adquirir información vinculante con el crimen ya que generalmente quedan rastros de imágenes, mensajes, consultas. En la materialización de la amenazas ya es tarde actuar, sin embargo, es lo que se intenta evitar, analizando proactivamente los mensajes con el fin de minimizar los riesgos y que estos lleven a cosas peores.

En el trabajo de Amato et. al. (2009) se propone un prototipo de aprendizaje de máquina para detectar imágenes con contenido adulto enviadas a través de mensajes *MMS* (*Multimedia Message Service*) en un celular con sistema operativo *Symbian*. El prototipo plantea interceptar la imagen, analizarla y clasificar el contenido de la imagen utilizando un servicio remoto. Este servicio utiliza un algoritmo de clasificación binario llamado *Support Vector Machine* (*SVN*), el cual está entrenado para identificar imágenes con contenido de carácter sexual. En este sentido, el aplicativo propuesto en el presente trabajo tiene mucho en común con el trabajo de Amato, para lo cual se estudiarán los mecanismos que utilizaron los autores para llegar a analizar las imágenes, clasificar el contenido (binario en el caso de Amato), para posteriormente realizar un entrenamiento de máquina y determinar si la imagen tiene contenido sexual o no.

Para la clasificación de textos en los mensajes de *WhatsApp* con el fin de determinar la existencia de un tipo de delito como SEXO, DROGAS, *BULLYING* en los menores de edad es necesario centrarse en los avances científicos sobre inteligencia artificial en el campo del análisis de lenguaje natural, en el cual se emplean un proceso de etiquetación de las palabras, Graña Gil, J. (2002), con la finalidad de clasificar las palabras en función del

léxico de un lenguaje en particular para determinar relaciones o conexiones entre las palabras, es así que se puede determinar un secuencia lógica o hilo conductor de por ejemplo una conversación, discurso, pregunta, tema, etc. Para lograr obtener una predicción de lenguaje natural, se observa que el algoritmo *Naíve Bayes* Multinomial o simplemente *MultinomialNB* es uno de los mejores y más utilizados algoritmos para realizar clasificación de texto y predicciones. Muchas aplicaciones de investigación relacionadas al análisis de textos por ejemplo se tiene: *Análisis de polaridad en textos escritos en inglés y español* de Plaza Sacarrera, L. (2014), *Visualización y seguimiento de acontecimientos en Twitter* de MARÍA, J., & GINER, M. (2017), *Aprendizaje activo para clasificación de preguntas* Teruel, M. (2015), *Predicción de acciones de control aéreo* de González Sendino, R. (2017).

Para la adquisición de la base de datos de *WhatsApp* seguiremos el procedimiento indicado por Bosschert (2017), quien explica el mecanismo para obtener la base de datos de *WhatsApp* desde el dispositivo móvil hacia una fuente externa. Este procedimiento es realizado mediante el desarrollo de una aplicación tipo *Android* que es instalada en el dispositivo móvil. Este procedimiento servirá para conocer sobre programación de aplicaciones en *Android* y la adquisición de la base de datos de *WhatsApp*.

Con respecto al conocimiento sobre los detalles técnicos de la base de datos de *WhatsApp* servirá la información publicada por Sahu (2014). En su artículo sobre análisis forense de *WhatsApp* en dispositivos móviles *Android*, da a conocer la ubicación de la base de datos y el sistema de archivos de un dispositivo *Android*. Además analiza la estructura de tablas y campos de la base de datos en cuestión. En este contexto, está información servirá para la adquisición de la base de datos de *WhatsApp* y para conocer su estructura.

El análisis para el acceso *root* a dispositivos móviles fue analizado por parte de Faheem et. al. (2014), en este artículo a más del acceso *root*, se realiza una descripción de la arquitectura de dispositivos *Android* (almacenamiento). Por otro lado, se describe una serie de comandos útiles para obtener el acceso de *root*[10] en un dispositivo *Android*. Este procedimiento de acceso *root* servirá para la adquisición de la base de datos de *WhatsApp*, la misma que se encuentra en un lugar del sistema de archivos de acceso restringido.

Una vez que se haya obtenido la base de datos, se necesita aplicar un procesamiento sobre los datos para el análisis y obtención de patrones mediante algoritmos. Para lograr esto, la solución planteada se apoyará en la librería *Scikit-learn*[11]. Librería que contiene herramientas para minería y análisis de datos. Además, es accesible y de libre descarga[12] e instalación por su licencia *Open Source BSD* ideal para proyectos en la academia. Por

[10] Previamente el dispositivo debe estar cambiado el sistema de arranque (flasheo).
[11] http://scikit-learn.org/stable/
[12] Para obtener Scikit Learn basta como ejecutar el comando pip install -U scikit-learn

último, esta librería implementa los mejores y más modernos algoritmos utilizados hoy en día para el análisis de datos mediante aprendizaje de máquina. Debido a esto, es necesario comprender el lenguaje *Python* ya que esta librería está implementada en este lenguaje de programación.

La constante innovación de las tecnologías trae entre otras cosas mejores resoluciones de pantalla, mayor rapidez en la carga y descarga de contenido en Internet, mejor experiencia de usuario en las aplicaciones que usamos, mayor conectividad entre grupos, amigos y familiares. Estos avances han venido acompañados de la aparición de mejores componentes de hardware, así como de la aparición de sistemas operativos más eficientes. Un sistema operativo exitoso a nivel mundial es *Android*, el mismo que fue lanzado el 23 de septiembre de 2008, lo que significó un auge en la utilización masiva de los *Smartphones*, acompañado de un alto contenido multimedia y una gran experiencia de usuario.

CAPÍTULO II

ESTADO DEL ARTE

2.1 Resumen

Este capítulo describe los trabajos de investigación relacionados con este proyecto. Primero se analiza los Aplicativos de Control Parental en Dispositivos Móviles, los cuales han ido evolucionando tanto desde el punto funcional como apoyo en la crianza del menor hacia sus padres.

Posteriormente, se describe los tipos y la clasificación de un aplicativo de control parental, integrando en un solo grupo las funcionalidades que estos ofrecen conforme a lo que existe actualmente en el mercado, y las investigaciones realizadas en este sentido.

Es importante poner énfasis en aplicativos que tienen como finalidad ofrecer servicio de monitoreo; los mismos que pueden servir para monitorear la navegación por internet, el uso de las aplicaciones, la descarga de aplicativo de pago o no, el tiempo que el menor dedica a su *Smartphone* y en especial aplicativos que monitorean mensajería y redes sociales. A este tipo de aplicativos móviles los denominaremos como aplicativos orientados al monitoreo.

Por último, se especifican trabajos de investigación orientados al análisis con NLP (Procesamiento de Lenguaje) en *WhatsApp*, los mismos que nos muestran las bondades que se puede lograr con este procesamiento y los mecanismos que han sido utilizados para lograr el objetivo de analizar y clasificar los textos.

2.2 Aplicativos de Control Parental en Dispositivos Móviles

Entre los principales estudios de software para control parental se ha encontrado el de Çankaya y Odabaş (2009), quienes analizan un tipo de control parental del tipo restrictivo que utilizan los padres frente al uso del computador, ya que, se identifican los riesgos y las afectaciones del uso extremo del computador: juegos con violencia, fácil acceso a pornografía y a contenido controversial. Incluso se detectó, mediante encuestas, que la privacidad en los adolescentes estaba comprometida, es así que se plantea y sugiere la responsabilidad y la toma de precauciones por parte de los padres proactivamente con la ayuda de un aplicativo que apoye en la labor del control parental.

Amato et. al. (2009) proponen un prototipo para detectar imágenes con contenido adulto enviadas a través de mensajes *MMS* (*Multimedia Message Service*) o por *Bluetooth* en un celular con sistema operativo *Symbian*. El sistema plantea interceptar la imagen, analizarla y clasificar el contenido de la imagen utilizando un servicio remoto. Este servicio utiliza un

algoritmo de clasificación binario llamado *Support Vector Machine* (*SVN*) el cual está entrenado para identificar imágenes con contenido de carácter sexual.

Por otro lado, en la Universidad Espíritu Santo de Guayaquil (*UESS*) se desarrolló un aplicativo para control parental (Larreátegui y Sánchez, 2016) quienes proponen realizar control parental a nivel del aplicativo con sistema Operativo *Android*, el propósito de este trabajo es apalear los efectos negativos en el uso extremo de los dispositivos inteligentes. El mecanismo que proponen para realizar el control parental, es la ejecución de un bloqueo en el dispositivo en horas o días de la semana previamente definidas por el padre, con el fin de evitar la distracción del menor por el uso del dispositivo. La arquitectura planteada se divide en 5 componentes: aplicación para padres, aplicación para hijos, servidor de aplicaciones, servidor de base de datos y servidor de mensajería *GCM* (*Google Cloud Messaging*).

Un prototipo de diseño de aplicativo de control parental elaborado por Hashish et. al. (2014) permite a los padres, conjuntamente con sus hijos, trabajar conjuntamente para seleccionar que tipo de aplicativos restringir de acuerdo a sus necesidades y conveniencias, a la vez de enseñarles qué aplicaciones son apropiadas y cuales no para sus hijos. En este trabajo se define los requisitos que debe cumplir un aplicativo de control parental y las funcionalidades que debe ofrecer. Por ejemplo, configurar restricciones y aplicar filtros con el fin de evitar contenido inapropiado para el menor.

Plan&Play es una aplicación basada en *Android* desarrollada por Hiniker, et. al. (2017). Esta aplicación aplica control parental a nivel de sistema, bloqueando aplicaciones y limitando el tiempo de uso en cada una de estas. Previamente el padre elabora un plan donde el aplicativo de control parental tiene previamente configurado que aplicaciones se usarán y el tiempo máximo de uso en cada aplicación. Además, permite presentar alertas en la pantalla al menor, con la finalidad de exhortar a dejar de usar un aplicativo en particular.

Preliminarmente, en base a estos trabajos se puede pre clasificar, de acuerdo a las funcionalidades que los aplicativos de control parental ofrecen en: monitorear, restringir, bloquear aplicaciones y su tiempo de uso, analizar el contenido de mensajería, analizar el uso de las aplicaciones, limitar el tiempo de uso de las aplicaciones, restringir compras por *Google Play*. Cada aplicativo tiene su particularidad y se diferencia de la adopción de una u otra opción según su mercado de competencia y la estrategia a seguir para diferenciarse de los demás aplicativos.

2.3 Clasificación de Aplicativos de Control Parental en *Android*

Para entender las funcionalidades que ofrecen los aplicativos de control parental y como estos se estructuran, se utilizará en un principio un estudio de Wisniewski et. al. (2017), quienes analizan un grupo de 75 aplicativos orientados a promover la seguridad en adolescentes. Este trabajo clasifica al grupo de aplicaciones seleccionados de control parental disponibles a la fecha del estudio, de esta forma identifica un grupo de 42 características principales que debe tener un aplicativo de este tipo. Además, proponen el *Conceptual TOSS Framework (Teen Online Safety Strategies)* el cual divide en dos aristas conceptuales el control parental: por un lado el Control Parental y por otro la Auto Regulación del menor acompañada de supervisión paternal. Dentro de este marco de trabajo identifican las categorías primarias del control parental: monitoreo, restricción y mediación activa; además, de tres categorías análogas a la estrategia de autorregulación adolescente: auto monitoreo, control de impulsos y afrontamiento de riesgos conforme los muestra la figura 2.

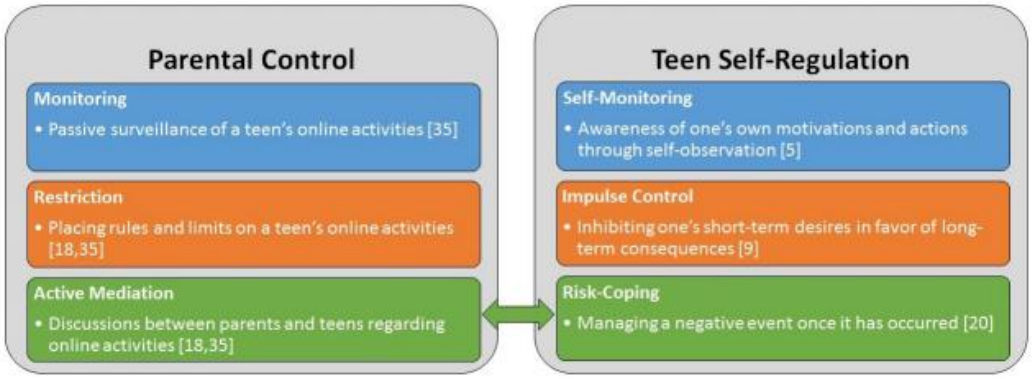

Figura 2. *Teen Online Safety Strategies (TOSS) de Wisniewski*, et al. (2017)

En este estudio se menciona la necesidad de crear aplicativos que no solo monitoreen y restrinjan, sino que además apoyen al control parental. Los autores sugieren "el objetivo es buscar mejores vías para aprovechar los principios del diseño de la participación colaborativa, la cual respete las necesidades y fomente los retos y oportunidades que son únicas no solo para los jóvenes, sino para los padres, en este sentido algunas investigaciones recientes han empezado a usar los métodos de diseño de participación para reducir el uso adictivo del *Smartphone* y el *Cyberbullying*, dando resultados positivos que conducen en esa dirección." (Wisniewski et.al., 2017)

Santos, et. al. (2013) plantean una clasificación de las características básicas más comunes que debe tener un aplicativo de control parental. Para esto, categorizan a este tipo de aplicativos en 5 grupos, y en cuyo trabajo proponen un aplicativo de control parental que se

compone de Lanzador de Aplicaciones, Llamadas, Mensajería, Notas y Configuraciones los cuales poseen propiedades de cada una de las categorías antes mencionadas.

1. **Lanzador (*Launcher*)**
   Bloquear la pantalla de inicio.
   Restringir instalación y ejecución de aplicaciones.
   Bloquear y descargar en la tienda de aplicaciones (*Play Store*).
   Bloquear cambios en la configuración del sistema.
   Bloquear accesos al sistema de archivos.
   Bloquear la navegación en Internet.
   Bloquear y restringir la aplicación de correos.

2. **Llamadas**
   Control de restricción de las llamadas.
   Permitir llamadas salientes a números autorizados.

3. **Mensajería**
   Restringir Enviar / Recibir mensajes de acuerdo a los contactos autorizados.

4. **Notas**
   Toma de notas.
   Mantener rastro de las notas.
   Recordatorios.

5. **Configuraciones (*Settings*)**
   Permitir configuraciones de otras aplicaciones.
   Lista de todas de las actividades desarrolladas en la plataforma.
   Registros de llamadas.
   Registros de *SMS*.
   Registros de notas.
   Grabación de fotos.
   Rastreo de la localización y los sitios que el menor visitó.

Otro trabajo en este campo se llama Sistema Remoto de Control Parental usando Smartphones, *RePort* por sus siglas en inglés (*Remote Parental Control System using Smartphones*), este es un modelo para control parental planteado por Kuppusamy et. al. (2013), conforme a la figura 3, el cual va dirigido a la supervisión o monitoreo de un *Smartphone*. En este modelo se plantea las siguientes acciones para un aplicativo de control parental: monitoreo, reportes, alertas y filtros.

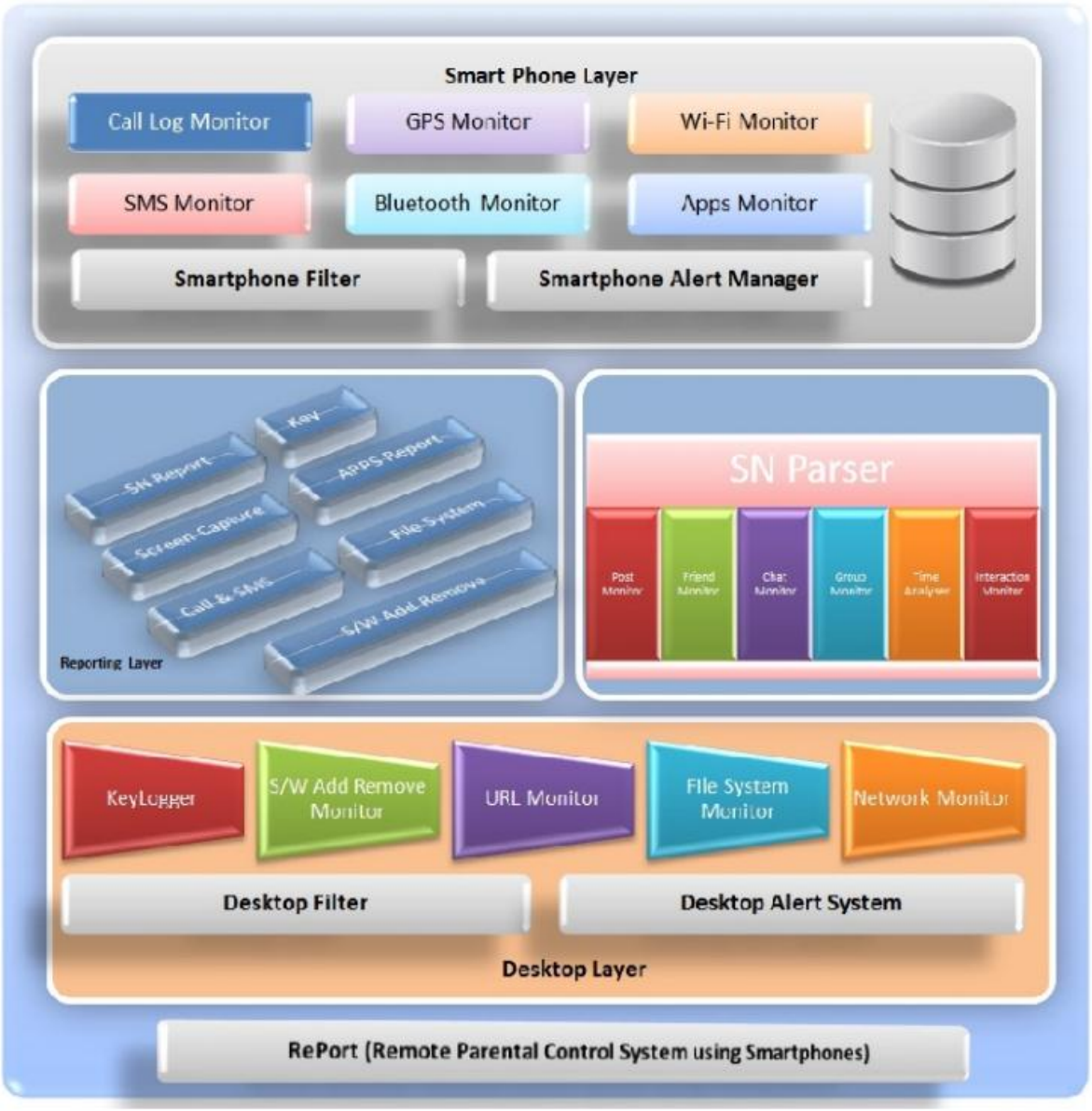

Figura 3. Modelo tomado de *RePort* de Kuppusamy et. al. (2013)

Las características que nos plantean estos tres trabajos sobre control parental (*Framework TOSS*, *RePort* y el trabajo de Santos et. al. (2013)), se las puede integrar en dos categorías principales según las funcionalidades que ofrecen las aplicaciones de control parental, esto es: monitoreo y restricción. Como se puede apreciar en detalle en la tabla 1.

Aplicativos de Control Parental en Dispositivos Móviles	
Monitoreo	**Restricción**
• Rastreo y localización de sitios visitados. • Actividades desarrolladas en la plataforma. • Monitorear sitios navegación. • Monitorear mensajería. • Monitorear contenido multimedia. • Monitorear *WI-FI, Bluetooth*	• Restringir instalación, ejecución aplicaciones. • Ejecución, Descargas, instalación de aplicaciones. • Configuración del sistema y de aplicaciones. • Accesos al sistema de archivos. • Navegación en Internet. • Llamadas entrantes / salientes a números autorizados. • Enviar / Recibir mensajería

Tabla 1. Resumen de Categorías en las Aplicaciones de Control Parental

Cabe indicar que, en cualquiera de estas dos categorías, algo en común que manejan los aplicativos de control parental es la facultad de enviar alertas hacia destinatarios interesados del control parental. El mecanismo de envío depende de cada aplicativo según la estrategia de diseño que se defina con la finalidad de alertar a los padres. Es importante hacer énfasis que el presente trabajo se centrará en investigar el monitoreo como metodología para el control parental.

2.4 Sub clasificación de Aplicativos de Control Parental que realizan Monitoreo

Dentro de la categoría de Monitoreo de Aplicaciones de Control Parental se realiza una sub clasificación basada en los trabajos sobre control parental de *Framework TOSS, RePort* y Santos et. al. (2013). Para este caso se denominará sub categorías, a aquellas que se encargan de tareas de monitoreo en el dispositivo móvil como:

- Localización – *GPS:*
  Obtiene la localización (latitud y longitud) del dispositivo, con el fin de ubicar geográficamente al menor en caso de emergencia.
- Redes de Comunicación:

Realiza monitoreo a los datos que son intercambiados por el *WIFI, modem, o Bluetooth* para descifrarlos y determinar actividades inusuales, informando a un agente externo sobre estas actividades.

- Actividades:
  Realiza monitoreo sobre actividades dentro del dispositivo móvil como, instalación y desinstalación de aplicaciones, uso de aplicaciones (tiempo de uso, frecuencia), con esto se puede enviar alertas o informes proactivas al padre.
- Navegación:
  Monitorea el uso de páginas de Internet, histórico de visitas, tiempo de visitas en las páginas, existen aplicativos que rastrean la navegación en el sistema de archivos del dispositivo.
- Mensajería:
  Analiza el contenido de los mensajes *SMS, MMS*, o *Chat* en redes sociales y el emisor de los mismos.
- Contenido Multimedia.
  Analiza el contenido de las imágenes, audios, videos que han sido descargados, o transferidos (*Bluetooth*) con la finalidad de detectar contenido malicioso o perjudicial.
- Mensajes del Sistema:
  Analiza los mensajes de log lanzados por el sistema o por las aplicaciones, con el fin de detectar actividades anómalas dentro del dispositivo como por ejemplo, fallas en los componentes del equipo, software malicioso en el dispositivo producto de instalaciones desde fuentes no oficiales a *Google Play*, análisis de *logs* para detección de instalación y desinstalación mal intencionada de aplicaciones y servicios por parte del menor o de terceros quienes pueden estar tratando de alterar o dañar el dispositivo. El padre puede conocer estas actividades oportunamente para un control parental proactivo.

Además del análisis de los trabajos antes mencionados, con lo que se obtuvo estas categorías, se realizó un sondeo de las aplicaciones de control parental con mayor descarga que se encuentran en *Google Play Store* con el fin de clasificarlos según las funcionalidades y opciones que ofrecen, las cuales se detallan a continuación:

- Localización:
  Permite el monitoreo para poder ubicar geográficamente al dispositivo en caso de robo o emergencia.
- Aplicaciones:
  Monitorean las aplicaciones y el uso que se da en cada una de ellas.
- Actividades:
  Realiza monitoreo sobre actividades como, instalación y desinstalación de paquetes, uso de aplicación (tiempo de uso, frecuencia).

- Internet/Historial:
  Monitorea el uso de páginas de Internet, histórico de visitas, tiempo de visitas en las páginas.
- Llamadas:
  Monitorean las llamadas realizadas por emisores que constan en la lista de contactos y de contactos desconocidos.
- Mensajería:
  Analizan la mensajería de *SMS, MMS* o mensajes en redes sociales.
- Monitoreo En Línea:
  Permite monitorear las actividades en el instante en que ocurren los eventos.
- Zona Segura:
  Permite mantener el dispositivo en estado de zona segura, como por ejemplo de navegación por internet evitando el ingreso a sitios no idóneos para menores, descarga e instalación de aplicaciones en el dispositivo y ejecución de aplicaciones con contenido peligroso.
- Contactos:
  Permite monitorear los contactos y alertar en el caso de llamadas de contactos desconocidos.
- Botón de Pánico:
  Si se mantiene presionada unos segundos, una tecla del dispositivo previamente definida (o conjunto de teclas), enviará una alerta de peligro mediante la plataforma del proveedor del aplicativo hacia un contacto en particular registrado en la instalación.
- Borrado Remoto:
  Permite el borrado remoto de aplicaciones que sean catalogadas como peligrosas o nocivas para el menor.
- Alertas:
  Permite el envío de alertas vía *SMS* o correo electrónico sobre eventos inusuales en los dispositivos.
- Procesamiento de Lenguaje Natural:
  Esta funcionalidad analiza los textos de los mensajes *SMS, MMS* o mensajes en redes sociales con el fin de determinar si el contenido está siendo peligroso.
- Descargas:
  Monitorea las descargas tanto las realizadas en el Internet como las que se realizan por medio del *Google Play*, para alertar sobre aplicaciones con contenido de adultos o que requieren algún tipo de pago.

En la tabla 2 se resume todas las funcionalidades antes detalladas, con cada una de las categorías encontradas en el *Framework TOSS, RePort* y *Santos et. al. (2013)*.

Monitoreo								
Sub Clasificación	Localización – GPS	Redes de Comunicación	Actividades	Navegación	Mensajería	Contenido Multimedia	Mensajes del Sistema	Servicios Adicionales
Funcionalidades	Localización Activa / Pasiva	WIFI	Llamadas	Sistema de archivos	SMS, MMS	Fotos	Sistema	Zona Segura
	Lugares visitados	Bluetooth	Contactos	En el Internet	Correo-e	Videos	Aplicaciones	Botón de pánico
		Módem	Instalación, desinstalación		Redes Sociales	Voz		Control Remoto
			Configuración del sistema		Procesamiento de Lenguaje Natural			Envío de Alertas
			Almacenamiento					

Tabla 2. Funcionalidades por sub categorías del tipo de Aplicativo Monitoreo

2.4.1 Aplicaciones de Monitoreo en el Google Play

Una vez obtenidas las funcionalidades y opciones necesarias para realizar un correcto monitoreo, se realizó una comparación de las aplicaciones con mayores descargas del *Google Play Store.* Esta comparativa se basa en las funcionalidades descritas en el apartado anterior, las cuales fueron encontradas gracias a la investigación de Santos et. al. (2013), el *Framework TOSS, RePort* y el sondeo de aplicaciones de control parental realizado en el *Google Play Store*. Este resultado se resume en la tabla 3.

Aplicación	A	B	C	D	E	F	G	H	I	J	K	L	M	N
UISEK Control Parental						X	X						X	
Control Parental SecureTeen	X	X	X	X								X		X
Kids Place - Control Parental y Bloqueo Infantil	X													
Qustodio		X												
SecureKids			X		X	X								
ESET Parental Control								X						
Control parental Screen Time		X												
Kaspersky SafeKids									X		X			
Control parental Safe Family		X			X	X	X							
Controles parentales Kids Zone									X					X
Parentsaround Parental control				X				X		X				
TeenLimit control parental	X					X								

Aplicación	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Sentry - Control Parental Inteligente				X										
Kid's Shell - Kids Corner								X		X				
Mi Sereno : Control Parental					X						X			
Parental Control Light				X										
Norton Family parental control						X								
Protect Your Kid											X			
FamilyTime Control parental														
Safe Browser Control Parental			X			X	X		X	X		X		
Control parental y localizador familiar	X													
Samsung Kids					X									
Screen Time Companion App			X											
Mobile Fence Parental Control								X		X				
Modo Niños - Bloqueo infantil					X		X							
McAfee Family Protection			X											
Safe Lagoon Control de Padres							X							
Control Paterno Funamo			X		X									
Kids Mode	X			X					X			X		

Aplicación	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Pumpic		X												
MobiStealth			X											
FootPrints						X								
mSpy				X										X
Bark						X	X						X	

Tabla 3. Funcionalidades de los aplicativos del *Google Play*

Leyenda			
A. Localización	E. Llamadas	I. Contactos	M. Procesamiento de Lenguaje Natural
B. Aplicaciones	F. Mensajería	J. Botón de Pánico	N. Descargas
C. Actividades	G. Monitoreo En Línea	K. Borrado Remoto	
D. Internet/Historial	H. Zona Segura	L. Alertas	

2.5 Procesamiento de Lenguaje Natural en *WhatsApp*

Debnath et.al. (2016) plantean la posibilidad de realizar minería de datos a los mensajes intercambiados de *WhatsApp* por los socorristas que hacen frente a las emergencias, esto debido a que las redes sociales poseen una gran capacidad de difusión y son grandes portadores de información. En este estudio se plantea la posibilidad de utilizar el conocimiento que se genera en las redes sociales para una respuesta efectiva antes los desastres. El objetivo del análisis de los mensajes de *WhatsApp* se da con el fin de suplir las necesidades que se requieran en el lugar de la emergencia en el menor tiempo posible.

En base a la referencia de Miller et. at. (2007), *WordNet* fue desarrollado por un grupo de psicólogos y lingüistas en 1985 en la Universidad de *Princeton* como un aporte para la investigación basado en el modelo taxonómico de Miller (1985). De acuerdo a Montraveta y Vázquez (2010), *WordNet* es una gran base de datos léxica, que ha sido diseñada para relacionar palabras y frases cercanas sintácticamente entre sí, la misma que contiene información relacionada a un lenguaje como, sinónimos, conceptos, hiponimia[13], y otras relaciones léxicas.

Para lograr el objetivo, se plantea realizar una minería de datos a los mensajes recolectados de los grupos de *WhatsApp*, mediante la utilización de técnicas de procesamiento de Lenguaje Natural, ayudado de técnicas de similitud semántica (en palabras y frases) y una base de datos léxico-estructurado llamado *WordNet*. Para esto, los investigadores utilizan los datos de los mensajes de *WhatsApp* para entrenar el modelo con los tópicos relacionados a una emergencia.

En el estudio de Debnath et.al. (2016), se concluye que sí los registros de los mensajes intercambiados pueden ser recolectados y analizados de un número grande de organizaciones encargadas de colaborar en los desastres y que trabajan al mismo tiempo y en el mismo lugar, puede ser útil para monitorear los requerimientos así como también la disponibilidad de diferentes recursos en tiempo real.

2.6 Discusión del estado de arte

Analizando los aplicativos disponibles en el mercado relacionados a la categoría de Monitoreo y que tienen la capacidad de analizar la mensajería, concentramos el análisis en características como tipos de redes sociales que analizan los aplicativos (p.e. *WhatsApp*), posibilidad de procesamiento de lenguaje natural y por último el idioma soportado, obteniendo como resultado la tabla 4 la misma que realiza una comparación entre las aplicaciones disponibles en el mercado versus las características definidas:

- Redes Sociales.
- Procesamiento de Lenguaje Natural.
- Idioma Soportado

[13] Hiponimia: Viene de hipónico que significa una palabra cuyo significado está contenido en el de otra.

Esta comparación en la tabla 4 se realiza debido a que la propuesta del aplicativo planteado se basa en el análisis de mensajería, en este caso de *WhatsApp*, ayudado de procesamiento de lenguaje natural para clasificar los textos. Este aspecto es un factor diferenciador entre los aplicativos de control parental analizados además que promueve más investigaciones al respecto, ya que se puede expandir las funcionalidades del aplicativo a otras redes sociales como *Facebook, Instagram*, etc. Además, que su análisis sintáctico es perfectible conforme al léxico de la región donde sea ejecutado, ya que no es el mismo lenguaje coloquial hablado por un mexicano, colombiano, español o ecuatoriano.

De entre las aplicaciones citadas en la tabla 4, la aplicación *Bark* es un fuerte competidor para la aplicación planteada en este proyecto, debido a que realiza el monitoreo de redes sociales utilizando procesamiento de lenguaje natural, sin embargo, este análisis de lenguaje natural se lo realiza en el idioma Inglés, lo que hace que carezca de soporte actualmente en países de habla hispana. Mientras que el aplicativo planteado en el presente trabajo procesa lenguaje español, con coloquialismos utilizados en el Ecuador.

Aplicación	**Redes Sociales**	**Procesamiento NLP**	**Idioma**
UISEK Control Parental	*WhatsApp*	X	Es (Ecuador)
mSpy	*WhatsApp*		
Qustodio	*Facebook, Twitter, Instagram, Whatsapp*		
Screen Time	*Facebook, Twitter, Instagram, Whatsapp*		
Norton Family	*SMS*		
FootPrints			

Aplicación	Redes Sociales	Procesamiento NLP	Idioma
Pumpic	*Instagram, Snapchat, Skype, WhatsApp, Kik, Viber, SMS y Facebook*		
MobiStealth	*Facebook, WhatsApp y Viber*		
Bark	*WhatsApp*	X	En (United Kingdom)

Tabla 4. Aplicaciones que monitorean redes sociales

CAPÍTULO III

SOLUCIÓN ADOPTADA

3.1 Introducción

La solución planteada se compone de una aplicación móvil para *Android*, que se encarga de obtener los mensajes de *WhatsApp* en el dispositivo. Un servicio que procesa, clasifica dichos mensajes y registra las alertas para luego ser enviadas a otro proceso que se encarga del pronóstico y categorización del mensaje; y un sitio *web* para consulta de los riegos a los que se expone el menor por parte del padre o madre que realizan el control parental.

3.2 Descripción de la Arquitectura

La aplicación sigue una arquitectura clásica cliente servidor, la cual permite la interconexión de los componentes. De acuerdo a la descripción de la figura 4, donde se puede divisar cada uno de los actores que forman parte de esta solución y la forma como cada uno de los componentes interactúan entre ellos. La arquitectura de la aplicación está compuesta por: Servidor Remoto, Aplicación *Web*, Aplicación Móvil.

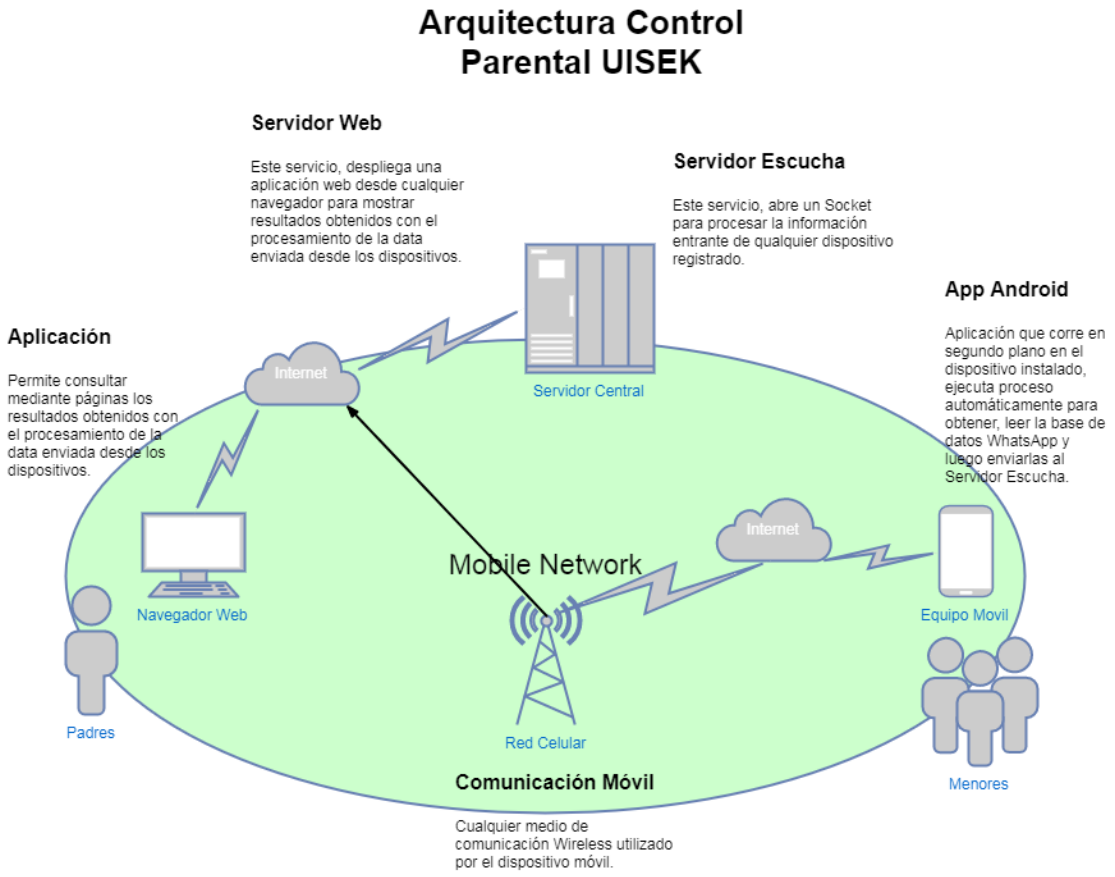

Figura 4. Diagrama de arquitectura de la solución

Como muestra la figura 4, la solución en primer lugar se compone de una aplicación *Android*, la cual se encarga de adquirir la base de datos de *WhatsApp*, leer los mensajes y enviarlos hacia un servidor remoto llamado Servidor *Socket*. Una vez que los datos arriban al servidor, son procesados y clasificados según el pronóstico que entregue el algoritmo y el clasificador *MultinomialNB* previamente entrenado con datos de prueba obtenidos a través del *API* de *Twitter*, con esto se logra el análisis, la clasificación y la identificación de un posible delito.

Posteriormente, el sistema envía una alerta al correo previamente especificado en la suscripción (durante el proceso de la instalación del aplicativo en el móvil) como apoyo para el control parental. Además, se dispone de un sitio web para consulta de los riegos a los que se expone el menor mediante un gráfico de barras, para esto se ingresa como filtro las fechas de inicio y fin, más adelante en este mismo capítulo se explica a detalle cada uno de los componentes.

Con la finalidad de describir cómo se relacionan los componentes de software y cómo interactúan entre ellos, el diagrama de componentes de la figura 5 permite describir los componentes y la forma como estos se comunican para lograr el funcionamiento global de la solución.

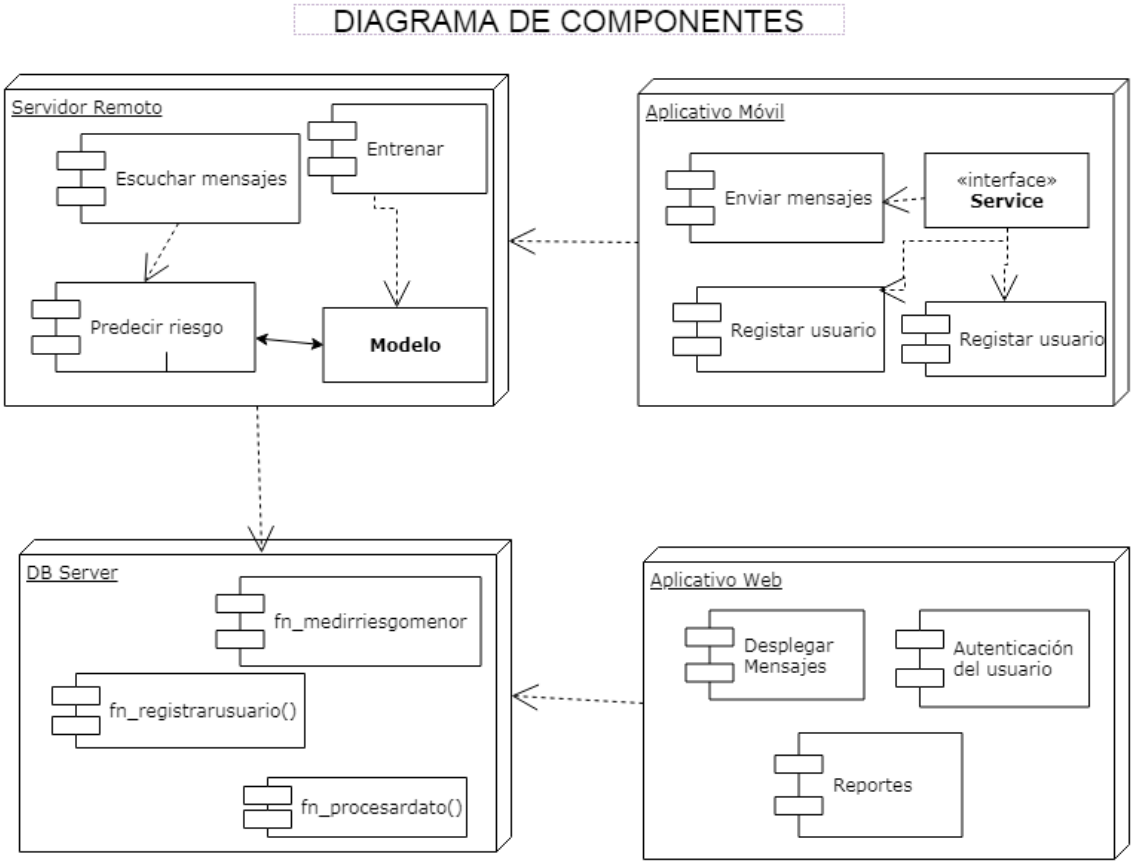

Figura 5. Diagrama de componentes de la solución

3.3 Software Base

El *software* base, con las respectivas versiones que requiere la solución, se detalla a continuación, este *software* servirá para el correcto funcionamiento y despliegue de la solución planteada desde el punto de vista de la aplicación móvil, Servidor Remoto, Aplicación Web, adicional a esto, se detalla también el *software* utilitario necesario para diferentes tareas como modelamiento de base de datos, editor de imágenes, etc.

Aplicativo Móvil

- *Android Studio 3.1 Canary 5*
  GUI para desarrollo, implementación y pruebas del proyecto.
- *Android SDK Tools version* 26, *Android* 4.1 (*Jelly Bean*)
  Herramientas y librería de interface para *Android* y pruebas.
- *Kingo SuperUser 1.2.0P*
  Aplicación para la administración de usuarios *root* en el aplicativo móvil.

Aplicativo Servidor Remoto

- *Python* 2.7
  Lenguaje de programación *Python*
- *Atom 1.25.1*
  IDE para desarrollo en *Python*
- *json_utils 0.2*
  Utilidad para procesar estructuras *JSON*
- *bson 0.5.0*
  Requerimiento de *Tweepy* y lectura de estructuras *JSON*
- *tweepy 3.5.0*
  Conexión y filtros de *Twitter*
- *pymongo 3.6.0*
  Estructuras anidadas tipo *JSON*
- *numpy 1.8.2*
  Procesamiento computacional científico para cálculos.
- *Scipy 0.13.3*
  Procesamiento matemático y gráficos estadísticos
- *scikit-learn 0.19.1*
  Librería para máquinas supervisadas, redes neuronales, algoritmos de clasificación, entrenamiento y regresión.

- tornado 3.1.1
  Librería de *Facebook* para gestión de *sockets* basado en eventos.

- *stop_words 2.23.1*
  Lista de palabras en varios idiomas incluido en español para filtrar palabras que no agregan valor en la transformación y entrenamiento.

- *psycopg2 3.0*
  Permite la inter conexión a bases de datos *PostgreSQL*

Aplicativo *Web*

- *Apache Server 2.2*
  Servidor *Web* para alojar el sitio y responder a las solicitudes *HTTP*

- *PostgreSQL 9.4*
  Almacena muestras de conversaciones categorizadas, suscriptores y mensajes de correo.

- *Bootstrap 4.0*
  Librería para diseño *Web responsive*.

- *adodb5 2.1*
  Librería para abstracción de datos.

- *jQuery 1.11.3*
  Librería *JavaScript* para interactuar con el *DOM*.

- *jQuery-ui 1.11.4*
  Librería Gráfica basada en *jQuery*.

- *Swiftmailer-5*
  Librería para el envío de correos.

- *CryptoJS 3.1.2*
  Librería para encriptar datos en el lado del cliente (*Javascript*), previo a enviar los datos por un canal inseguro.

Utilitarios

- *DB Browser for SQLite versión 3.10.1.*
- *PgAdmin 1.2.*
- *MySql Workbench 6.2 CE.*
- *Adobe Fireworks CS6.*

3.4 Modelo

La solución de software propuesta está compuesta por tres componentes principales, cada uno de ellos cumple una actividad específica lo cual permite finalmente alertar a los padres sobre eventos de peligro contra los menores. En base a esto, a continuación

se detalla cada uno de estos componentes con la finalidad de dar al lector un panorama claro de la solución completa.

3.4.1 Datos Base de Entrenamiento

El principal objetivo del aplicativo es el envío de alertas tempranas cuando se detecta la presencia de algún tipo de delito contra el menor de edad, el cual es el usuario del dispositivo móvil y que habitualmente envía y recibe mensajes a través del aplicativo de mensajería instantánea de *WhatsApp*. Bajo este contexto, como punto de partida se ha categorizado tres grupos de amenazas de riesgo contra los menores:

- Drogas
- Sexo
- *Bullying*

Basados en la clasificación de estos riesgos, se plantea la necesidad de contar con datos de entrenamiento, como insumo básico para el pronóstico. Con este propósito se toman mensajes que se generan en la red social *Twitter* ayudados de su *API* para este proceso.

Para obtener los datos se creó varios programas en el lenguaje Python, los cuales filtran ciertos tipos de mensajes de Twitter según sea la necesidad de filtro del riesgo o amenaza. En este caso se realizó filtros para *bullying*, sexo y drogas, además se agregaron mensajes del tipo neutral, es decir, mensajes que no genera ningún tipo de peligro para el menor y que contrasta con los otros filtros para dar un valor neutral.

En la tabla 5 se muestra un ejemplo de los *scripts* en *Python* preparados para la clasificación y un ejemplo del "*CORPUS*" utilizado para el entrenamiento y para el filtro para los datos de la máquina de predicción, cabe indicar que las palabras citadas son una muestra sacada del léxico de asihablamos.com[14] y fepe55.com[15] (código de la droga). Sin embargo, se pueden seguir refinando el listado, e incorporando nuevas palabras con respecto a los coloquialismos en el Ecuador para perfeccionar la clasificación con el fin que este quede libre de errores y así evitar malas interpretaciones.

Sexo maquinaaprender_sexo.py	**Bullying** maquinaaprender_bulling.py	**Drogas** maquinaaprender_drogas.py	**Neutral** maquinaaprender_neutral.py
topics = ['Puta', 'Culo', 'Chucha', 'Pene', 'Verga',	topics = ['Maricón', 'Perra', 'Zorra', 'Guey', 'Flaco',	topics = ['Droga', 'Cocaína', 'Mariguana', 'Heroína', 'La H',	topics = ['Paz', 'Amor', 'Felicidad', 'Amigo', 'Amistad',

[14] http://www.asihablamos.com/word/pais/EC/

[15] http://www.fepe55.com.ar/blog/2004/01/02/diccionario-los-codigos-de-la-droga/

Sexo maquinaaprender_sexo.py	**Bullying** maquinaaprender_bulling.py	**Drogas** maquinaaprender_drogas.py	**Neutral** maquinaaprender_neutral.py
'Culito', 'Ano', 'Clítoris', 'Arecha', 'Arecho', 'Semen', 'Sémen', 'Vagina', 'Zorra', 'Perra', 'Perro', 'Culiar', 'Culeo', 'Sexo', 'Tirar', … .. .	'Gordo', 'Hediondo', 'Apestoso', 'Patón', 'Marrano', 'Cerdo', 'Puerco', 'Greñuda', 'Putito', 'Putita', 'Aflojar', 'Aniñado', 'Loco', 'Demente', 'Atarzanar', … .. .	'Bate', 'Bazuco', 'Buco', 'Canuto', 'Yerba', 'Porro', 'Lote', 'Merca', 'Limado', 'LSD', 'Mota', 'Ganlla', 'Narco', 'Brujo', 'Billuzo', … .. .	'Paseo', 'Juntos', 'Papa', 'Mama', 'Querer', 'Te quiero', 'Te amo', 'Leer', 'Lectura', 'Estudiar', 'Estudioso', 'Pana', 'Dios', 'Bendecir', 'Bendiga', … .. .

Tabla 5. Ejemplo de Palabras para los filtros según la categoría definida

Es importante indicar que la clasificación NEUTRAL nos dará resultados positivos cuando una conversación no contiene contenido nocivo o perturbador para el menor, con esto logramos evitar que el sistema clasifique de forma errónea una conversación y lance alertas falsas.

Para realizar la filtración de los datos, mediante una cuenta de *Twitter* previamente creada, se configura una aplicación en *Apps de Twitter*[16]. Mediante esta configuración se podrá acceder para filtrar los *Tweets* desde otro programa, en este caso desde la máquina de aprendizaje que describiremos en este mismo apartado, *Twitter* nos entrega cuatro valores de seguridad de acceso a la cuenta para el acceso a los *Tweets*, *access_token, access_token_scecret, consumer_key, consumer_secret*.

Con el acceso a *Twitter* y previamente definidas las categorías y los datos, se procedió a crear los *scripts* que realizan el filtrado. Estos están descritos en la tabla 6, cada uno tiene su particularidad según el contenido que se desea filtrar.

[16] https://apps.twitter.com/

Script	Tipo
maquinaaprender_bulling.py	Python file
maquinaaprender_drogas.py	Python file
maquinaaprender_neutral.py	Python file
maquinaaprender_sexo.py	Python file

Tabla 6. Vista de los *scripts* que recolectan datos del *Twitter*

La estructura de directorios mostrada en la figura 7, donde se almacenan los mensajes procesados se puede visualizar en la figura 6, donde cada archivo representa un mensaje en *Twitter* y es nombrado con la Fecha/Hora en formato numérico del momento en el cual es recolectado por el programa preparado para este proceso, al abrir un archivo al azar se puede observar el contenido.

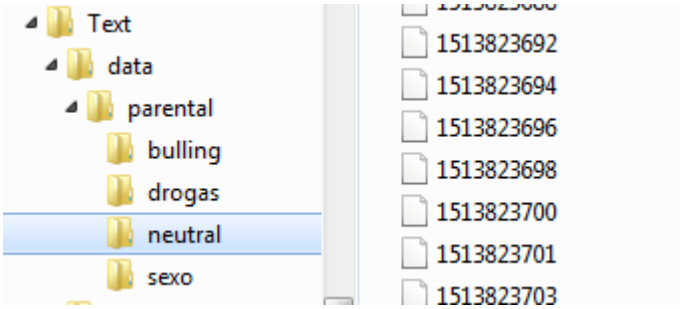

Figura 6. Estructura de carpetas y los documentos de los *tweets*

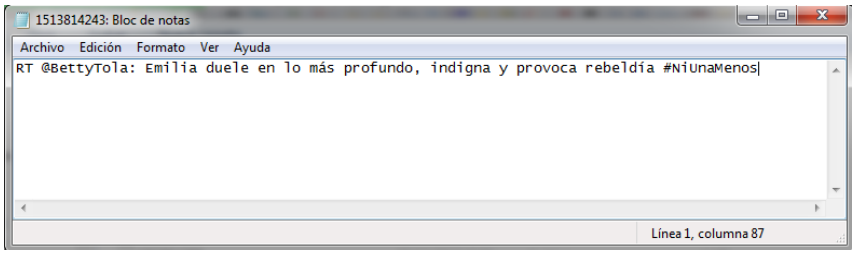

Figura 7. Ejemplo de un documento almacenado y su *corpus*

3.4.2 Entrenamiento de Máquina

Para realizar la clasificación de los mensajes, nos hemos basado en el ejemplo de la herramienta *Scikit*[17] llamado *document_classification_20newsgroups*, el cual utiliza entradas de un blog sobre tópicos *'alt.atheism'*, *'talk.religion.misc'*, *'comp.graphics'*,

[17] http://scikit-learn.org/stable/auto_examples/text/document_classification_20newsgroups.html

'sci.space', entre otros 20 grupos previamente clasificados. En el ejemplo del sitio scikit-learn se realiza un pronóstico, con la finalidad de identificar a qué tópico o tema pertenece un tipo de discusión. Basado en este ejemplo, se adaptó el programa para nuestro objetivo propuesto, es decir, una conversación se clasifica conforme a los tópicos delineados 'drogas', 'sexo', 'bullying', 'neutral', el algoritmo probabilístico clasificador *MultinomialNB()* o Multinomial *Naíve Bayes* y lo que la función *clf.predict()* retorne. El *script Pronosticar.py* es capaz de realizar este pronóstico por los datos y el entrenamiento previamente realizados.

En este caso, la clase *Algoritmo.py* (ver tabla 11), procesa el directorio que almacena todos los datos (mensajes) de ejemplo, obtenidos de *Twitter* y previamente clasificarlos de acuerdo a su contenido. Luego de la clasificación, los documentos son categorizados y transformados para ayudar al pronóstico. Este proceso, genera un archivo *cache* llamado *controlparental.pkz*. Este paso se realiza con el fin de ejecutar el pronóstico de manera rápida sin tener que procesar todo nuevamente reduciendo procesamiento de *CPU*, consumo de memoria y obteniendo un tiempo de respuesta bajo ya que no tiene que procesar todos los archivos cada vez que se hace la consulta.

3.4.3 Modelo de base de datos

3.4.3.1 Tablas

El modelo de base de datos se compone de tres tablas principales conforme a lo descrito por la figura 8. Estas tablas son:

1. **Tabla correo.-** Almacena los correos que el sistema genera, por ejemplo correo de registro de suscripciones, envío de alertas, etc.

Campo	Tipo	Resumen
id	int	Campo numérico representa la clave primaria de la tabla
de	character varying(50)	Almacena el correo del remitente
para	character varying(50)	Almacena el correo del destinatario
copia	character varying(50)	Almacena un correo copia del destinatario
contenido	text	Almacena el correo en formato *HTML*

Campo	Tipo	Resumen
fecha	timestamp	Almacena la fecha de generación del correo
estado	character(1)	Almacena el Estado del registro I = Ingresado, E = Enviado
imei	character varying(50)	Clave foránea que almacena el *IMEI* relacionado con el *ID* del dispositivo, se usa como clave foránea relacionada con el dispositivo.
asunto	character varying(200)	Almacena el asunto del correo

Tabla 7. Detalle de los campos de la tabla correo

2. **Tabla de usuarios.-** Almacena los usuarios suscritos al servicio cada vez que el aplicativo ha sido instalado.

Campo	Tipo	Resumen
imei	character varying(50)	Campo *ID* primario de la tabla que almacena el *ID* del dispositivo llamado *IMEI*
nombres	character varying(100)	Almacena los nombres del suscriptor
apellidos	character varying(100)	Almacena los apellidos del suscriptor
clave	character varying(100)	Almacena la clave del suscriptor con encriptación *SHA256* para seguridad
correo	character varying(200)	Almacena el correo del suscriptor el cual recibirá las alertas
actual	int	Almacena el valor de riesgo actual, con este valor el sistema determinará si es tiempo de enviar la alerta o no, se

Campo	Tipo	Resumen
		basa en el valor del campo nivel-riesgo
Correo copia	character varying(200)	Correo de copia para el envío de las alertas, es configurable desde el aplicativo *Web*
nivel-riesgo	numeric	Nivel de riesgo que tolerará el padre medido en Porcentaje 0% ~ 100%

Tabla 8. Detalle de los campos de la tabla usuarios

3. **Tabla datopronostico.-** Almacena los datos que llegan al servicio de clasificación, en estos datos se guarda el pronóstico entregado por el algoritmo en el componente Servidor Remoto. Cuando un nuevo registro es almacenado, se ejecuta un *trigger*.

Campo	**Tipo**	**Resumen**
id	int	*ID* numérico que representa la clave primaria de la tabla
fecha	timestamp	Fecha en que se registró el dato de la conversación
dato	text	Dato contexto de la conversación mantenida el cual fue enviado por el dispositivo
pronostico	character varying(50)	Valor del pronóstico realizado por el algoritmo de clasificación, sus valores pueden ser DROGAS, SEXO, *BULLYING*, NEUTRAL
nivel	smallint	Nivel de riesgo del mensaje, sirve para cálculos en los envíos de alertas, sus valores posibles son 0 (inexistencia de riesgo) o 5 (existencia de riesgo)

Campo	Tipo	Resumen
imei	character varying(50)	*ID* de clave foránea del dispositivo que envió el dato

Tabla 9. Detalle de los campos de la tabla datopronostico

La relación entre cada una de las tablas es mediante el *ID* de dispositivo, el cual es almacenado en el campo *IMEI*, este *ID* viene acompañado en cada dato que llega al servidor. Este valor es un identificador único asignado al dispositivo por el fabricante y es así que se puede diferenciar entre varias conversaciones de varios suscriptores y estas se las puede relacionar con las otras tablas del modelo.

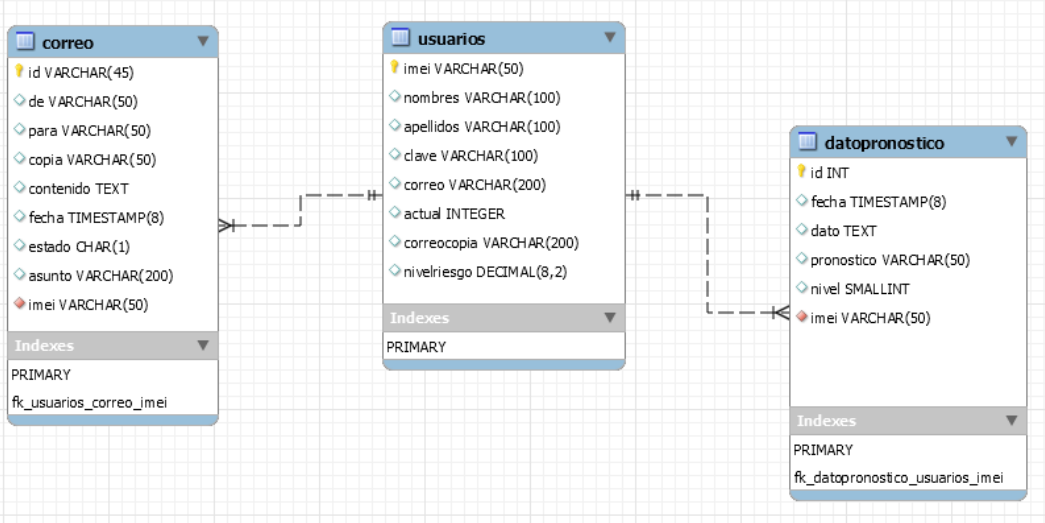

Figura 8. Diagrama del modelo de base de datos del aplicativo

3.4.3.2 Funciones y Desencadenantes (*Triggers*)

Existen objetos del tipo función y *triggers* que realizan tareas dentro de la base de datos con la finalidad de apoyar al aplicativo en los procesos de registro de eventos de mensajería para su posterior envío.

Objeto	Tipo	Definición	Resumen
fn_procesardato	Función	fn_procesardato(imei text, cadena text, pronostico text,	Procesa y registra el dato que ha sido invocado desde el Servidor Remoto

Objeto	Tipo	Definición	Resumen
		ip text, puerto integer)	
fn_registrarusuario	Función	fn_registrarusuario(_imei text, _nombres text, _apellidos text, _correo text)	Función para registrar una nueva suscripción de usuario desde el Servidor Remoto
fn_medirriesgomenor	Función (Trigger)	fn_medirriesgomenor()	Función que se invoca mediante un *trigger* y que se encarga de determinar el nivel de riesgo que existe para el menor, y si es el caso registra un correo de alerta en la tabla del modelo llamada *correo* que será enviada desde el *script correo.php*
tgrmedirriesgomenor	Trigger	tgrmedirriesgomenor AFTER INSERT ON datopronostico FOR EACH ROW EXECUTE PROCEDURE fn_medirriesgomenor();	Ejecuta la función fn_medirriesgomenor cada vez que un registro es almacenado en la tabla del modelo *datopronostico*

Tabla 10. Funciones y *triggers* que intervienen en el registro de los datos

3.4.4 Componente - Aplicativo móvil

Este componente se encarga de obtener los mensajes de *WhatsApp* (como muestra el diagrama de la figura 9). Es decir, realiza la lectura de los datos, para luego, mediante la

inter-conexión de *socket* al servidor remoto de la plataforma que ha sido dotado de una *IP* pública, envía dichos datos disponibles en ese instante.

Cabe indicar que este componente se instala a manera de servicio en el dispositivo móvil, lo que lo hace invisible para el usuario, es decir, se ejecuta en segundo plano sin la necesidad de intervención o interacción con el usuario. En el caso de reinicio o apagado del dispositivo, el servicio vuelve a ser iniciado automáticamente por el mismo sistema operativo.

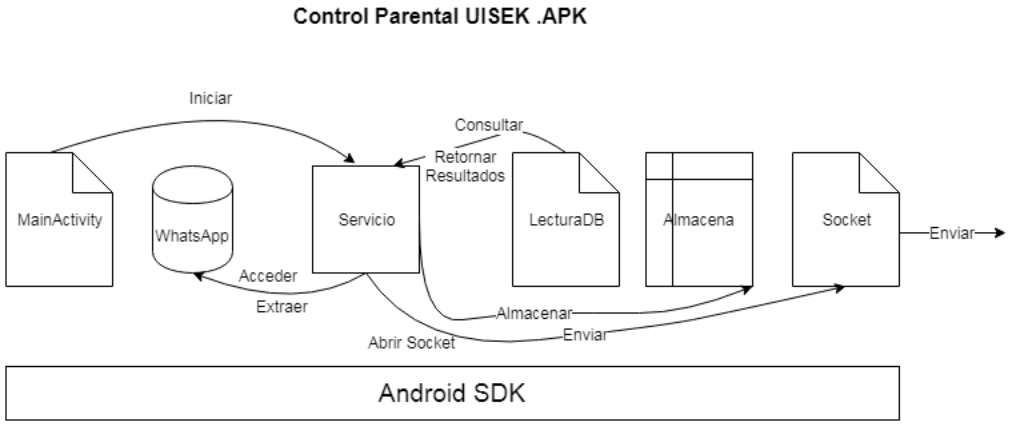

Figura 9. Diagrama de clases del aplicativo parental

Este servicio ejecuta un proceso de extracción y lectura de los mensajes de *WhatsApp* de manera periódica. Está diseñado para siempre leer los mensajes recientes, a través del registro del último dato enviado.

Una parte importante del programa es la extracción de datos, cuyo proceso reside en el método `EnviarDatosSocket.ExtraerBaseDatosWA`, el cual ejecuta el comando *dd* basado en sistemas *UNIX* para copiar los mensajes hacia una carpeta definida para uso del aplicativo.

En detalle el proceso de extracción de la base de datos de *WhatsApp*, consiste en la ejecución de los siguientes comandos dentro del aplicativo:

2. Obtener el acceso *root* con el comando su
3. Dar permiso de propiedad a la base de datos *msgstore.db* con el comando
4. `chmod 664 /data/data/com.whatsapp/databases/msgstore.db`
5. Copiar desde la carpeta origen hacia la carpeta destino deseada:
6. `dd if=/data/data/com.whatsapp/databases/msgstore.db of=%s/ControlParental/msgstore.db`

Es importante indicar que estos pasos son posibles gracias a que el equipo móvil, ha sido previamente *flasheado* y *rooteado*, lo cual no es otra cosa que el proceso de instalar

un *software* personalizado diferente al que viene de fábrica y luego con la ayuda de un programa de administración se logra otorgar permisos de *root* al aplicativo.

3.4.5 Componente - Servidor remoto

Para procesar los mensajes de las conversaciones enviados desde el dispositivo móvil y hacia el servidor remoto se ha diseñado y creado un servicio interno, el cual abre un *socket* para atender cada comunicación establecida entre un dispositivo móvil y el Servidor Remoto. Este componente se encarga de recibir las conversaciones desde los dispositivos móviles, para luego procesarlas y analizarlas con el fin de determinar si en el mensaje existió algún tipo de riesgo en su contenido en contra el menor en el contexto de DROGAS, SEXO, *BULLYING* o NEUTRAL (el tipo NEUTRAL significa que no ha existido ningún tipo de delito).

El componente servidor remoto está diseñado para estar habilitado a manera de servicio en el lado del servidor, su función principal es de recibir los datos que el aplicativo móvil envía mediante una interconexión *socket* en el puerto *1234* (número de puerto definido para el aplicativo). Cuando detecta una conexión abre un puerto y espera la transferencia de los datos que envía el servicio instalado en el dispositivo móvil, el *socket* se cierra una vez hecha la transferencia de datos, a continuación, los mensajes son almacenados y finalmente categorizados con la ayuda de un algoritmo de clasificación previamente entrenado. Acotando, el componente servidor remoto realiza un procesamiento interno programado es el encargado de generar y enviar un correo de confirmación al suscriptor registrado por el aplicativo durante el proceso de instalación. El flujo en el cual interactúan los elementos del Servidor Remoto se encuentra en la tabla 11 y en el diagrama de la figura 10.

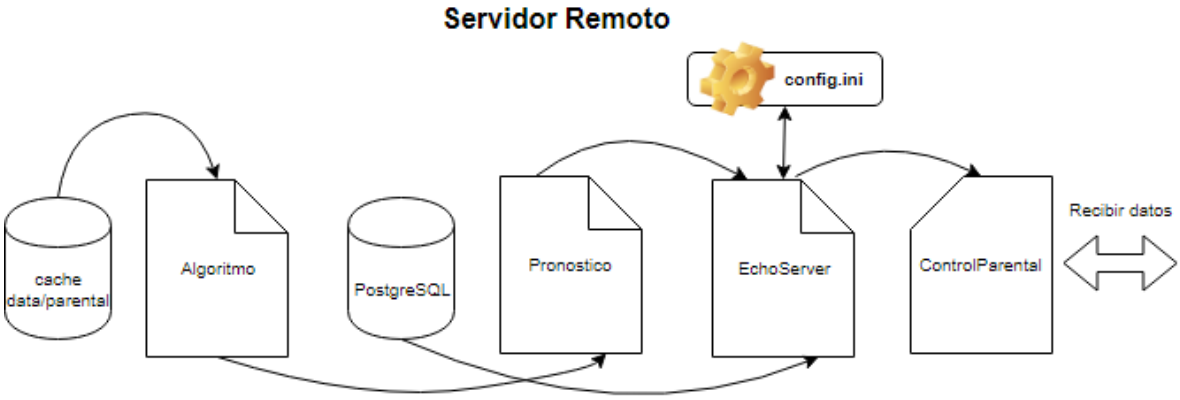

Figura 10. Proceso de interacción en el Servidor remoto

Explicando más a detalle la figura 10, el proceso en el servidor remoto inicia al ejecutar el *script ControlParental.py* el cual contiene la función principal *main()* del aplicativo. Esta función ejecuta las sentencias *_server = EchoServer()* y *_server.listen('192.168.100.2')* que son funciones del *script ServidorEscucha.py*. Este *script* al momento de ser instanciado, en primer lugar ubica y lee los parámetros del archivo de

configuración *config.ini*, para consecuentemente abrir un *socket* en el puerto *1234* (parámetro configurable que lo lee del archivo *config.ini*). Con esto, el servidor está listo para escuchar mensajes que envían los dispositivos.

Cuando llega un mensaje, se ejecuta el *script Pronostico.py*, este *script* realiza dos tareas, la primera es invocar al *script Algoritmo.py* que lee el almacén de datos recolectados en función de las categorías definidas y pasadas como parámetro (SEXO, DROGAS, *BULLYING*, NEUTRAL). La segunda tarea consiste en invocar al clasificador *MultinomialNB* para que entregue un pronóstico en base a la entrada. Una vez invocado este algoritmo se retorna un pronóstico, el mismo que es almacenado en la base de datos mediante le ejecución de la función *fn_procesardato()*, cuya explicación se revisó en la sección *3.4.3 Modelo de base de datos*.

Para revisar a detalle la composición de carpetas y archivos en directorio de el Servidor Remoto, se procede a explicar detalladamente cada uno de los *scripts* que forman parte de este proceso de servidor. La explicación del registro y envío de los mensajes de alerta se enuentran en las secciones *3.4.6 Componente - Aplicativo Web y 3.4.4 Modelo de base de datos*.

Elemento	Tipo	Resumen
Data	Directorio	Carpeta que contiene los *tweets* recolectados, almacenados y clasificados según las categorías SEXO, DROGAS y *BULLYING*
Logs	Directorio	Almacena los archivos de los mensajes sobre errores de la aplicación
Algoritmo.py	*Script*	Contiene el mecanismo que procesa la información y genera un archivo *cache* con la extensión *pkz* para una ejecución más eficiente
config.ini	Archivo de configuración	Archivo que contiene configuraciones útiles para el aplicativo, como usuario de base, nombre de base de datos, puertos de conexión y las direcciones *IP*
EchoServer.py	*Script*	Script que realiza la apertura de un *socket* para comunicación con los dispositivos de forma remota y posteriormente procesa los mensajes recibidos.

Elemento	Tipo	Resumen
Pronosticar.py	*Script*	*Script* que apoyado con el *Algoritmo.py* ejecuta el pronóstico a determinar.
ControlParental.py	*Script*	Contenedor de la función principal *main()*, crea una instancia de *EchoServer* para iniciar el *socket*.

Tabla 11. *Scripts* que componen el Servidor Remoto

3.4.6 Componente - Aplicativo *Web*

La solución cuenta con una interfaz *Web* descrita en la tabla 12. Esta aplicación *web* tiene como objetivo apoyar a los padres. Para esto, mediante la interfaz *Web* se puede conocer a detalle los mensajes y alertas enviadas por el servidor. Esta interfaz web ayuda de manera gráfica y tabular a la comprensión y descripción de la información de alertas enviadas hacia los padres, de esta forma sirve como herramienta de apoyo para un mejor control parental.

Titulo	Tipo	Resumen
Images	Directorio	Almacena las imágenes de la aplicación
Libs	Directorio	Almacena las librerías que utiliza la aplicación para su normal funcionamiento
config.php	*Script*	Página para customizaciones dentro de la aplicación
conexión.php	*Script*	Almacena los parámetros de configuración tanto del aplicativo como para la base de datos
correo.php	*Script*	*Script* que procesa los correos que genera el aplicativo para ser enviados mediante la ejecución de un cron configurado en el sistema operativo tipo *Linux*
datos.php	*Script*	*Script* que retorna una estructura del tipo *JSON* que sirven para desplegar el reporte gráfico sobre las alertas enviadas

Titulo	Tipo	Resumen
index.php	*Script*	Página de inicio del aplicativo después del inicio de sesión
login.html	*Script*	Página de presentación para el ingreso de credenciales y autenticación por parte del usuario
mensaje.php	*Script*	Despliega el detalle sobre un mensaje o correo enviado por el aplicativo
mensajes.php	*Script*	Muestra la lista de mensajes históricos enviados a la cuenta del padre que está realizando el control parental.
validar.php	*Script*	*Script* que valida si las credenciales enviadas de la página *login.html* son válidas.

Tabla 12. Estructura del proyecto *web*

Para el lanzamiento de alertas, la aplicación web cuenta con un *script* llamado *correo.php* (descrito en la tabla 12). Este *script* cumple la función de consultar a la tabla *correo* de la base de datos la existencia de nuevos mensajes. En la sección *3.4.3 Modelo de la Base de Datos,* se detalla como los objetos de la base de datos realizan actividades para poblar la tabla correo con mensajes del aplicativo.

Para que el *script correo.php* sea ejecutado constantemente se lo configura como un proceso de servidor *CRON*[18], con este se asegura que deba ejecutarse cada minuto con el fin de detectar si hay alertas nuevas de correo (registradas en la tabla *correo* de la base de datos) para posteriormente enviarlas por correo electrónico al suscriptor.

Para actividades de envío de correos de prueba se ha creado una cuenta en el *Gmail* de Google, la cual ha sido denominada como *controlparentaluisek@gmail.com*, los correos son enviados desde el *Script correo.php* a través de la cuenta de *Gmail*.

Un correo informativo, como el del ejemplo desplegado en la figura 11, detalla el tipo de delito identificado por la aplicación. Un correo de este tipo es enviado por el proceso de alertas, con el fin de comunicar al padre de familia sobre la posible presencia de un delito en los mensajes de *WhatsApp* del menor.

[18] *CRON* proceso automático en sistemas operativos que se encargan de ejecutar *scripts* cada cierto tiempo

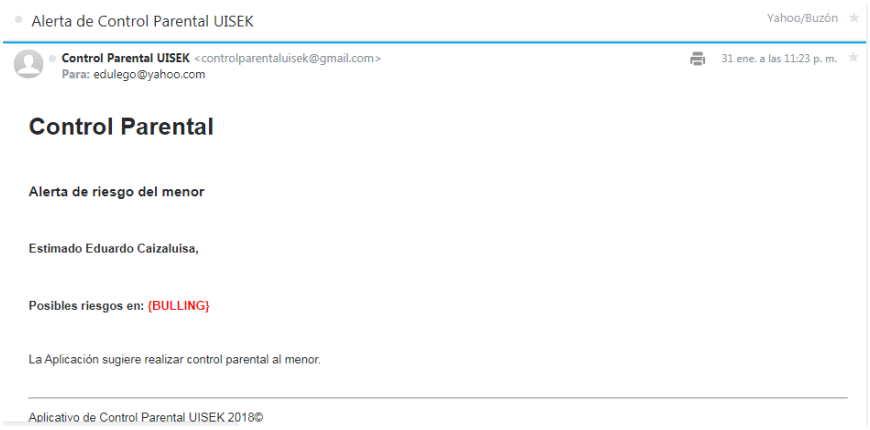

Figura 11. Correo de alerta del aplicativo de control parental

3.4.7 Interacción entre componentes

Los componentes indicados en la figura 5, y detallados en las secciones *3.4.3 Modelo de base de datos, 3.4.4 Componente - Aplicativo móvil, 3.4.5 Componente - Servidor remoto, y 3.4.6 Componente - Aplicativo Web*, interactúan entre ellos como se indica en la figura 12, la cual explica el proceso de interacción con el modelo de datos y la relación con los demás componentes del aplicativo.

El proceso de almacenamiento inicia cuando el Servidor Remoto recibe los mensajes de los dispositivos y ejecuta el pronosticador para obtener un pronóstico. Una vez obtenido esto, ejecuta la función *fn_procesardato()* la cual envía los datos hacia la base de datos que internamente almacena los datos en la tabla *datopronostico*. Atado a la tabla datopronostico está un *trigger* que ejecuta la función *fn_medirriesgomenor()*, la misma que se encarga de medir el riesgo en el mensaje en algunas de estas categorías SEXO, DROGAS, *BULLYING*, para conocer más a detalle del proceso de generación y envío de alertas, dirigirse a la *sección 4.3* Proceso de envío de alertas.

La forma como el Servidor *Web* interactúa con el modelo es a través de la generación de consultas a la tabla *datopronostico* para desplegar reportes. La tabla correo para enviar correos y para presentar un reporte con los históricos de los correos, la tabla usuarios para la autenticación del usuario y para almacenar personalizaciones de la cuenta de usuario.

Cabe destacar, que al ser una aplicación que se puede ejecutar de forma distribuida, cada componente puede ser instalado en un servidor independiente.

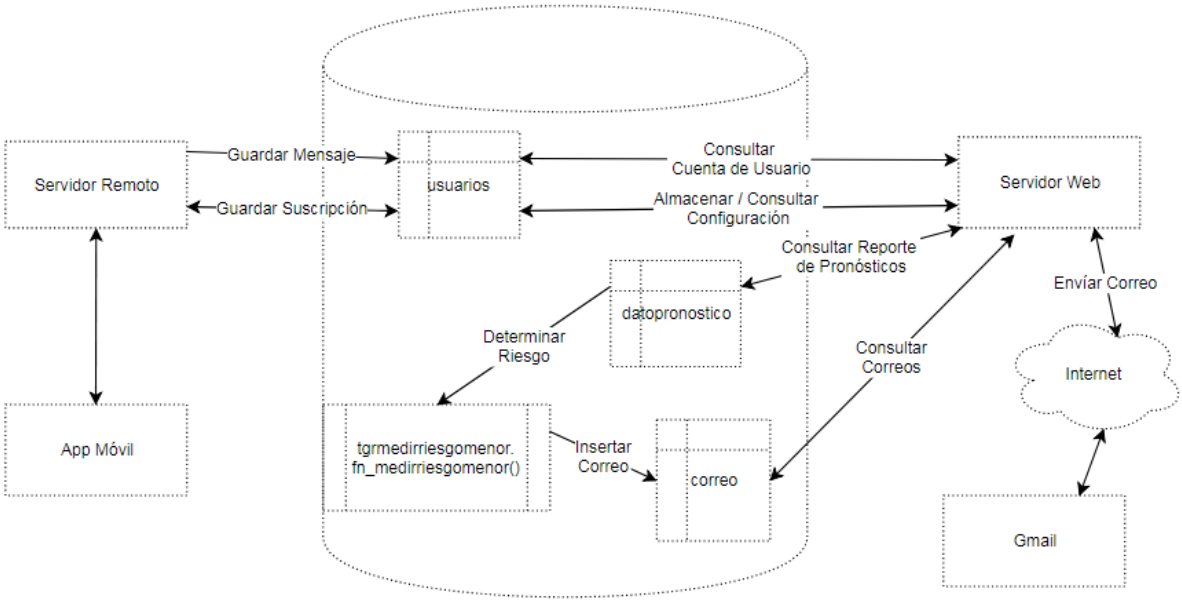

Figura 12. Flujo de interacción entre Componentes del Aplicativo y objetos del modelo

3.4.8 Instalación y despliegue del aplicativo

Al realizar la instalación del *APK* en un dispositivo móvil, el sistema solicita al usuario confirmar el permiso de acceso *ROOT* al aplicativo como se imprime en la pantalla de la figura 13. Este permiso es necesario para que la aplicación tenga acceso a la base de datos de *WhatsApp*. Para esto es necesario seleccionar "PERMITIR".

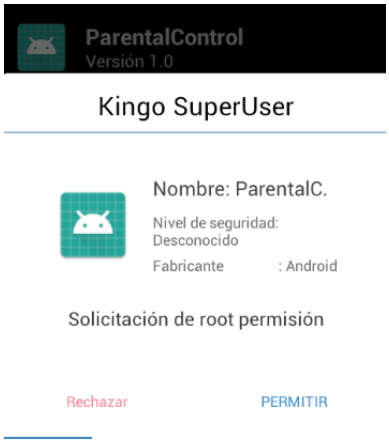

Figura 13. Solicitud de permisos *root* durante la instalación del aplicativo

Después de realizada la instalación del aplicativo para fines didácticos, se puede verificar que el aplicativo reside en la lista de aplicaciones del sistema operativo *Android* del dispositivo utilizado para pruebas. Se puede observar en la figura 14 como se despliega en la lista de Aplicaciones el proceso *ParentalControl*.

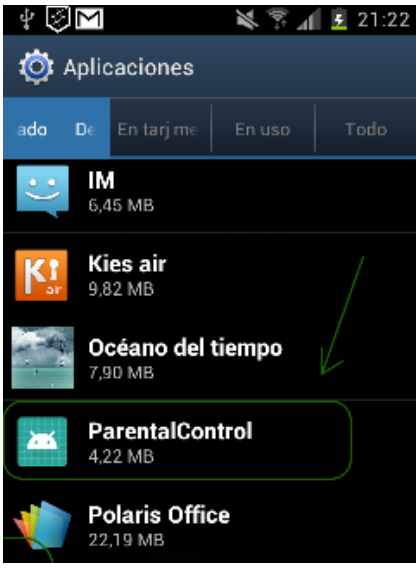

Figura 14. Lista de aplicativos del móvil donde se lista el aplicativo de Control Parental

Se puede revisar el detalle del aplicativo dentro del conjunto de aplicaciones en ejecución y se observa que el aplicativo está corriendo y tiene características de residir como servicio del sistema operativo (explicado en la figura 15). Además, este servicio será iniciado automáticamente pese a que el usuario apague o reinicie el terminal.

Figura 15. Vista del Servicio residente en el dispositivo móvil

Una vez que el aplicativo se instala correctamente en el dispositivo, se despliega un formulario de suscripción para ingresar los datos de la persona que va a realizar el monitoreo de control parental en el aplicativo *web* (ver la figura 22). Esta persona será quien reciba las alertas de control parental al correo que registre en este formulario. Una vez realizado el registro, un correo de confirmación será enviado, el mismo contendrá las credenciales de acceso al aplicativo *web*.

Consecuentemente, estos datos son enviados al servidor remoto donde son procesados y almacenados. Una vez registrados los datos de suscripción, se envía un correo con la confirmación. En este correo se puede encontrar datos para el acceso al aplicativo *web*, así como también la clave asignada a la cuenta, la misma que es generada automáticamente por el sistema.

Figura 16. Formulario de registro presentado después de la instalación

En resumen, el usuario tiene 4 pasos para completar el proceso de instalación de acuerdo a lo descrito en la figura 17.

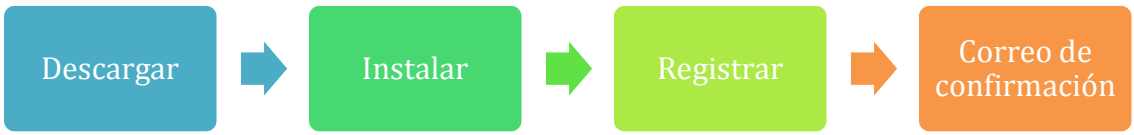

Figura 17. Pasos del proceso de instalación del aplicativo

CAPÍTULO IV

VALIDACIÓN EMPÍRICA

4.1 Método

El método a aplicar es la validación empírica, por medio de la ejecución de pruebas con conversaciones simuladas. Esto quiere decir, que dado una conversación de entrada se analizará los resultados obtenidos por la aplicación propuesta, con el fin de determinar la validez de la misma. Este proceso se detalla en el diagrama de la figura 18.

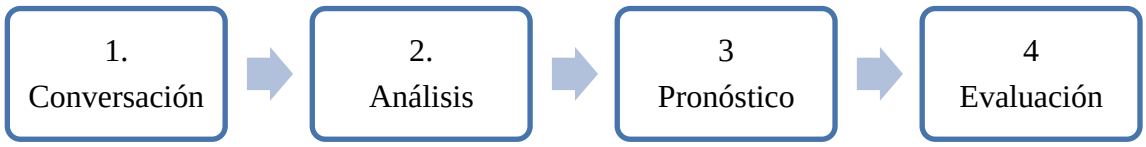

Figura 18. Pasos para el método de validación de las pruebas

La validación consiste en emular conversaciones en *WhatsApp*, las cuales serán recogidas por la aplicación móvil, para después ser analizadas por el servidor remoto.

El pronóstico entregado por el servidor será evaluado comparando la temática del mensaje. Por ejemplo. se enviarán frases como “Longo puerco de mierda”, el cual debe ser clasificado (por criterio del autor) como *bullying*, que luego de ser analizado por el aplicativo, como se puede observar el resultado en la figura 19.

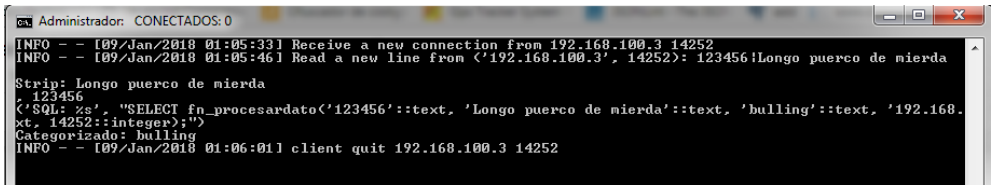

Figura 19. Resultado del envío de una frase obscena

Otro ejemplo, en este caso, al escribir una frase con obscenidades sobre sexo, el programa pronostica “*sexo*”, como se indica en la figura 20.

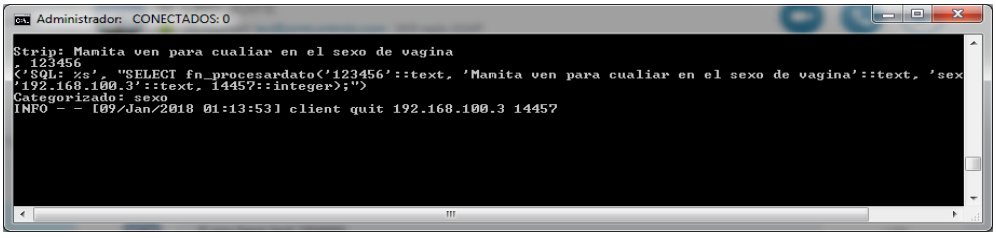

Figura 20. Resultado de escribir una frase relacionada con SEXO

4.2 Simulación de las conversaciones

Para la ejecución de prueba se ha preparado 4 conversaciones en función de las categorías previamente definidas es decir SEXO, DROGAS, *BULLYING* y NEUTRAL, teniendo así que el resultado esperado deberá enmarcarse en la categoría en donde se ejecute la conversación.

Estas pruebas siguen el flujo que indica la figura 21, la cual se inicia cuando el emisor envía un mensaje hacia el receptor. El aplicativo de control parental (instalado en el receptor) verifica la existencia de nuevos mensajes y de ser positiva dicha comprobación, envía cada mensaje de uno en uno hacia el servidor remoto, en donde el sistema registra y determina de qué tipo de riesgo se trata (SEXO, DROGA, *BULLYING*), finalmente un evaluador toma nota del resultado que el sistema entregue.

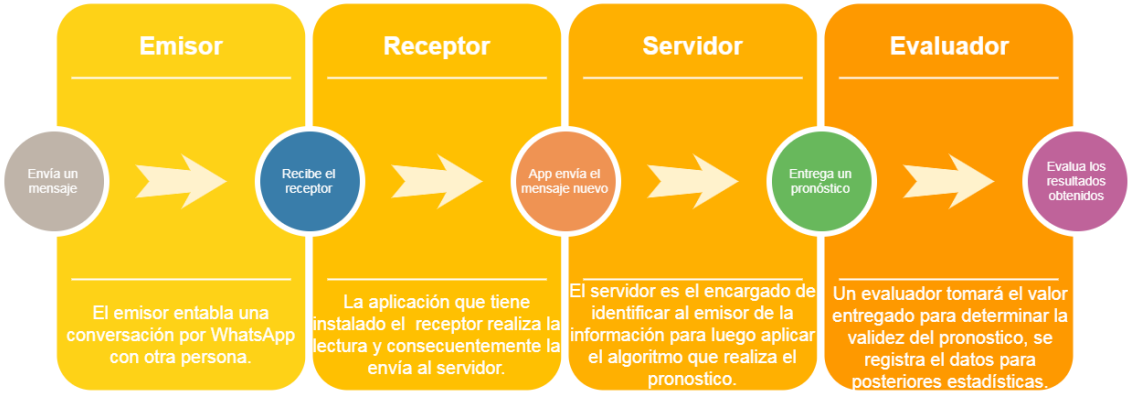

Figura 21. Flujo de la simulación de las pruebas

4.3 Resultados

Para obtener y evaluar los resultados, se plantea la ejecución de 2 intentos de validación, el primero se realizará sobre un banco de datos recolectados de *Twitter*, en donde cada categoría tiene un número de documentos diferente: *Bullying*, 16307; Drogas, 3885; Neutral, 42610; Sexo, 18693. La razón de la diferencia en la cantidad de documentos, se debe a que, en el momento de la recolección de los documentos (*Tweets*) una categoría recoge más documentos que la otra en el mismo rango de tiempo.

En el segundo intento de validación, se balancea la cantidad de documentos recolectados, haciendo que todas las categorías tengan la misma cantidad de documentos (66477) por cada categoría.

Con este paso, se evalúan a continuación los resultados entre el primer y el segundo intento en dos escenarios diferentes.

4.3.1 Primer Intento de Validación

En el primer intento de validación, se toma en cuenta el número de documentos recolectados por el *API* de *Twitter* a través de los *scripts* preparados previamente (ver en la *sección 3.4.1*). Para este intento se considera un desbalance en la cantidad de documentos con los que cuenta cada categoría, como se puede apreciar en la tabla 13, siendo la categoría Neutral con 42610 documentos la más alta y la categoría Drogas con 3885 documentos la más baja.

CATEGORIA	DOCUMENTOS
Bullying	16307
Drogas	3885
Neutral	42610
Sexo	18693

Tabla 13. Número de documentos recolectados en cada categoría

Con este antecedente, las pruebas en cada una de las categorías se presentan a continuación en las tablas 14, 15, 16 y 17.

Caso de Prueba: *Bullying*

En la tabla 14 se muestra el resultado en la ejecución de pruebas para el caso *BULLYING*.

CONVERSACIÓN TIPO *BULLYING*			
Actor	**Contexto**	**Resultado**	**Cumple Si/No**
Persona1	Longo puerco de mierda	*BULLYING*	SI
Persona2	Eres un estúpido tu	*BULLYING*	SI
Persona1	Callate tu no sirves para nada	*BULLYING*	SI

Persona2	Tienes caca en la cabeza	*BULLYING*	SI
Persona1	Te veo y te saco la mierda puto	*BULLYING*	SI
Persona2	Puto cagon te vas a arrepentir de decir esto	*BULLYING*	SI
Persona1	Meco marico le estas hablando a tu marido	*BULLYING*	SI
Persona2	Afeminado eres una niñita llorona	*BULLYING*	SI
Persona1	Sarnozo ladilla cállate	*BULLYING*	SI
Persona2	Mocoso deja de ser menestra	*BULLYING*	SI
Persona1	Jajaja	NEUTRAL	SI
Persona2	Adios cuídate o te mato	NEUTRAL	NO
		Total aciertos	**11/12**

Tabla 14. Diálogo de una conversación en el contexto *BULLYING*

Sacando los promedios para el caso de *BULLYING* obtenemos el porcentaje de 91.66% de efectividad.

Caso de Prueba: Drogas

En la tabla 15 se muestra el resultado en la ejecución de pruebas para el caso DROGAS.

CONVERSACIÓN TIPO DROGAS			
Actor	**Contexto**	**Resultado**	**Cumple Si/No**
Persona1	Fue una chupiza	NEUTRAL	NO
Persona2	No queria soltar el Norton	NEUTRAL	NO
Persona1	Simon cerveza	NEUTRAL	NO
Persona2	Bebi full alcohol despues de la fiesta	*BULLYING*	NO
Persona1	La pau tenia una tella	NEUTRAL	NO
Persona2	En la misma tella	NEUTRAL	NO
Persona1	Habia mezclado full norton	SEXO	NO

Persona2	La pau estaba hecho gato	NEUTRAL	NO
Persona1	Puta vamos a trickear	NEUTRAL	NO
Persona2	Armemos un bate fuera del cole	NEUTRAL	NO
Persona1	Heroina y extasis en la fiesta de amanda	NEUTRAL	NO
Persona2	El brujo me va a dar merca bien buena	NEUTRAL	NO
Persona1	Vendo mariguana en paquetes	NEUTRAL	NO
Persona2	Alcohol	DROGAS	SI
Persona1	Ya sabe	NEUTRAL	NO
Persona2	Full gansha	SEXO	NO
Persona1	Alcohol mijo	NEUTRAL	NO
Persona2	No te olvides sexo drogas y rockanroll	NEUTRAL	NO
Persona1	Para que armes un paquete de gansha	NEUTRAL	NO
Persona2	Tengo marihuana en paquetes	NEUTRAL	NO
Persona1	Es la grifa de lo mejor ya sabe	NEUTRAL	NO
Persona2	Estoy voladazo con esas drogas	NEUTRAL	NO
Persona1	Buena oportunidad para ir a trickear	NEUTRAL	NO
Persona2	Armemos un bate fuera del cole	NEUTRAL	NO
Persona1	Heroina y extasis lo que quiera papa	NEUTRAL	NO
Persona2	Habra alcohol?	NEUTRAL	NO
Persona1	Listo para la fiesta	*BULLYING*	NO
		Total aciertos	**1/27**

Tabla 15. Diálogo de una conversación en el contexto DROGAS

El resultado de la prueba DROGAS, en este caso al tabular obtenemos un promedio de 3.70% de efectividad.

Caso de Prueba: Sexo

En la tabla 16 se muestra el resultado en la ejecución de pruebas para el caso SEXO.

CONVERSACIÓN TIPO SEXO			
Actor	**Contexto**	**Resultado**	**Cumple Si/No**
Persona1	Hasta que tengas un orgasmo	NEUTRAL	NO
Persona2	Me gusta culiarte	NEUTRAL	NO
Persona1	Mija tienes el mejor culo	NEUTRAL	NO
Persona2	Eres un loco culion	*BULLYING*	NO
Persona1	Y terminamos en un motel	SEXO	SI
Persona1	Primero te chupo la vagina	SEXO	SI
Persona2	Te dare verga hasta el amanecer	NEUTRAL	NO
Persona1	Y me terminas dentro	NEUTRAL	NO
Persona2	Verga para ti	NEUTRAL	NO

Persona1	Y el semen adentro	SEXO	SI
Persona2	Quiero una vagina	SEXO	SI
Persona1	Tu culito rico	SEXO	SI
		Total de aciertos	**5/12**

Tabla 16. Diálogo de una conversación en el contexto SEXO

Tabulando y promediando el caso de SEXO se tiene el resultado de 41.66% de efectividad.

Caso de Prueba: Neutral

En la tabla 17 se muestra el resultado en la ejecución de pruebas para el caso SEXO.

CONVERSACIÓN TIPO NEUTRAL			
Actor	**Contexto**	**Resultado**	**Cumple Si/No**
Persona1	Los chicos estaran esperando	NEUTRAL	SI
Persona2	Mejor afuera	NEUTRAL	SI
Persona1	En la cafetería del cole o afuera	NEUTRAL	SI
Persona2	Prefiero de vainilla	NEUTRAL	SI
Persona1	Mañana comeremos un helado	NEUTRAL	SI
Persona2	Tu mama se va a enojar	NEUTRAL	SI
Persona1	Ya mejor andate a dormir	NEUTRAL	SI
Persona2	Mañana ser un mejor dia	NEUTRAL	SI
Persona1	Ya casi es hora de dormir	NEUTRAL	SI
Persona2	Mandame por correo	NEUTRAL	SI
Persona1	Si lo tengo	NEUTRAL	SI
Persona2	Tienes el deber de clase de naturales	NEUTRAL	SI
		Total aciertos	**12/12**

Tabla 17. Diálogo de una conversación en el contexto NEUTRAL

En el caso del tipo NEUTRAL tenemos una efectividad promedio perfecta 100%.

4.3.2 Segundo Intento de Validación

En el segundo intento de pruebas, se balancean los documentos recolectados por el *API* de *Twitter* para que todas las categorías manejen la misma cantidad de documentos, como indica la tabla 18. Con esto se desea comprobar si existe una mejora de los pronósticos con respecto a los resultados obtenidos en la sección *4.2.1 primer intento de validación*.

CATEGORIA	DOCUMENTOS
Bullying	66477
Drogas	66477
Neutral	66477
Sexo	66477

Tabla 18. Número de documentos en cada categoría

Caso de Prueba: *Bullying*

En la tabla 19 se muestra el resultado en la ejecución de pruebas para el caso *BULLYING*.

CONVERSACIÓN TIPO BULLYING			
Actor	**Contexto**	**Resultado**	**Cumple Si/No**
Persona1	Longo puerco de mierda	*BULLYING*	SI
Persona2	Eres un estúpido tu	*BULLYING*	SI
Persona1	Callate tu no sirves para nada	*BULLYING*	SI
Persona2	Tienes caca en la cabeza	*BULLYING*	SI
Persona1	Te veo y te saco la mierda puto	*BULLYING*	SI
Persona2	Puto cagon te vas a arrepentir de decir esto	*BULLYING*	SI
Persona1	Meco marico le estas hablando a tu marido	*BULLYING*	SI
Persona2	Afeminado eres una niñita llorona	*BULLYING*	SI

CONVERSACIÓN TIPO BULLYING			
Persona1	Sarnozo ladilla callate	*BULLYING*	SI
Persona2	Mocoso deja de ser menestra	*BULLYING*	SI
Persona1	Jajaja	NEUTRAL	SI
Persona2	Adios cuídate o te mato	*BULLYING*	SI
		Total aciertos	**12/12**

Tabla 19. Diálogo de una conversación en el contexto *BULLYING*

Sacando los promedios para el caso de *BULLYING* obtenemos el porcentaje de 100% de efectividad.

Caso de Prueba: Drogas

En la tabla 20 se muestra el resultado en la ejecución de pruebas para el caso DROGAS.

CONVERSACIÓN TIPO DROGAS			
Actor	**Contexto**	**Resultado**	**Cumple Si/No**
Persona1	Fue una chupiza	*BULLYING*	NO
Persona2	No queria soltar el norton	*BULLYING*	NO
Persona1	Simon cerveza	*BULLYING*	NO
Persona2	Bebi full alcohol despues de la fiesta	*BULLYING*	NO
Persona1	La pau tenia una tella	*BULLYING*	NO
Persona2	En la misma tella	*BULLYING*	NO
Persona1	Habia mezclado full norton	*BULLYING*	NO
Persona2	La pau estaba hecho gato	*BULLYING*	NO
Persona1	Puta vamos a trickear	*BULLYING*	NO
Persona2	Armemos un bate fuera del cole	*BULLYING*	NO
Persona1	Heroina y extasis en la fiesta de amanda	*BULLYING*	NO
Persona2	El brujo me va a dar merca bien buena	*BULLYING*	NO
Persona1	Vendo mariguana en paquetes	*BULLYING*	NO
Persona2	Alcohol	*BULLYING*	NO
Persona1	Ya sabe	NEUTRAL	SI
Persona2	Full gansha	*BULLYING*	NO
Persona1	Alcohol mijo	*BULLYING*	NO

CONVERSACIÓN TIPO DROGAS			
Persona2	No te olvides sexo drogas y rockanroll	*BULLYING*	NO
Persona1	Para que armes un paquete de gansha	*BULLYING*	NO
Persona2	Tengo marihuana en paquetes	*BULLYING*	NO
Persona1	Es la grifa de lo mejor ya sabe	*BULLYING*	NO
Persona2	Estoy voladazo con esas drogas	*BULLYING*	NO
Persona1	Buena oportunidad para ir a trickear	*BULLYING*	NO
Persona2	Armemos un bate fuera del cole	*BULLYING*	NO
Persona1	Heroina y extasis lo que quiera papa	*BULLYING*	NO
Persona2	Habra alcohol?	*BULLYING*	NO
Persona1	Listo para la fiesta	*BULLYING*	NO
		Total aciertos	**1/27**

Tabla 20. Diálogo de una conversación en el contexto DROGAS

El resultado de la prueba DROGAS, en este caso al tabular obtenemos un promedio de 3.70% de efectividad.

Caso de Prueba: Sexo

En la tabla 21 se muestra el resultado en la ejecución de pruebas para el caso SEXO.

CONVERSACIÓN TIPO SEXO			
Actor	**Contexto**	**Resultado**	**Cumple Si/No**
Persona1	Hasta que tengas un orgasmo	*BULLYING*	NO
Persona2	Me gusta culiarte	*BULLYING*	NO
Persona1	Mija tienes el mejor culo	*BULLYING*	NO
Persona2	Eres un loco culion	*BULLYING*	NO
Persona1	Y terminamos en un motel	*BULLYING*	NO
Persona1	Primero te chupo la vagina	*BULLYING*	NO
Persona2	Te dare verga hasta el amanecer	*BULLYING*	NO
Persona1	Y me terminas dentro	*BULLYING*	NO
Persona2	Verga para ti	*BULLYING*	NO
Persona1	Y el semen adentro	*BULLYING*	NO
Persona2	Quiero una vagina	*BULLYING*	NO
Persona1	Tu culito rico	*BULLYING*	NO
		Total de aciertos	**0/12**

Tabla 21. Diálogo de una conversación en el contexto SEXO

Tabulando y promediando el caso de SEXO se tiene el resultado de 0% de efectividad.

Caso de Prueba: Neutral

En la tabla 22 se muestra el resultado en la ejecución de pruebas para el caso NEUTRAL.

CONVERSACIÓN TIPO NEUTRAL			
Actor	**Contexto**	**Resultado**	**Cumple Si/No**
Persona1	Los chicos estaran esperando	*BULLYING*	NO
Persona2	Mejor afuera	*BULLYING*	NO
Persona1	En la cafetería del cole o afuera	*BULLYING*	NO
Persona2	Prefiero de vainilla	*BULLYING*	NO
Persona1	Mañana comeremos un helado	*BULLYING*	NO
Persona2	Tu mama se va a enojar	*BULLYING*	NO
Persona1	Ya mejor andate a dormir	*BULLYING*	NO
Persona2	Mañana ser un mejor dia	*BULLYING*	NO
Persona1	Ya casi es hora de dormir	*BULLYING*	NO
Persona2	Mandame por correo	*BULLYING*	NO
Persona1	Si lo tengo	*BULLYING*	NO
Persona2	Tienes el deber de clase de naturales	*BULLYING*	NO
		Total aciertos	**0/100**

Tabla 22. Diálogo de una conversación en el contexto NEUTRAL

En el caso del tipo NEUTRAL tenemos una efectividad de validación promedio de 0%.

4.4 Proceso de envío de alertas

Para determinar el envío de alertas, el sistema realiza un cálculo sobre el promedio diario de la cantidad de mensajes enviados en fechas anteriores, con esto se determina un valor base con el cual se decide si se envía o no de la alerta.

Por otro lado, el sistema se basa en un valor de nivel del riesgo para sacar un porcentaje, el cual se utiliza para configurar un umbral desde el cual se enviará la alerta.

Es decir, si el promedio de mensajes diarios enviados y recibidos es de 100 mensajes, si se selecciona un porcentaje de riesgo de 12% (como se muestra en la figura 22), el sistema lanzará la alerta al detectar que se alcanzó los 12 mensajes de riesgo sobre cualquiera de las categorías SEXO, DROGAS, *BULLYING*. Esto indica que el porcentaje de riesgo es directamente proporcional a los mensajes de riesgo enviados. Como otro ejemplo, si el nivel de riesgo está puesto al 1%, significa que de 100 mensajes enviados o recibidos, basta con un solo mensaje con riesgo para que el correo de alerta sea enviado por el sistema.

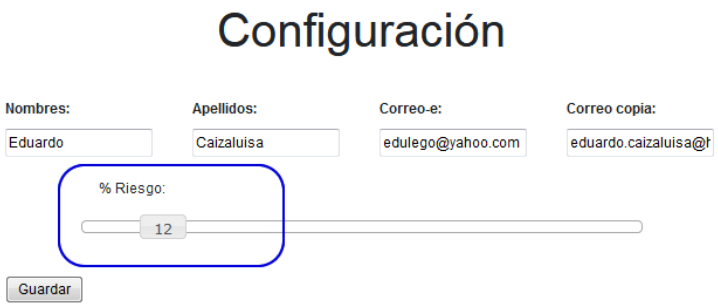

Figura 22. Configuración del aplicativo parental porcentaje de mensajes de riesgo.

Esta configuración es de gran ayuda a los padres si se desea dar un nivel de tolerancia bajo a los mensajes que recibe el menor.

Ejemplos de correos disparados por el sistema en la ejecución de las pruebas del primer intento de validación, se puede apreciar en las Figuras 23, 24 y 25.

Por ejemplo, para la expresión “Longo puerco de mierda”, la cual correspondería a la categoría *BULLYING*, la aplicación genera una alerta como la que se muestra en la figura 23.

Para la expresión “Alcohol”, la cual correspondería a la categoría DROGAS, la aplicación genera una alerta como la que se muestra en la figura 24.

Para la expresión “Primero te chupo la vagina”, que corresponde a la categoría SEXO, se genera la alerta por parte de la aplicación como se muestra en la figura 25.

Figura 23. Captura del correo de alerta lanzado por el aplicativo por *BULLYING*

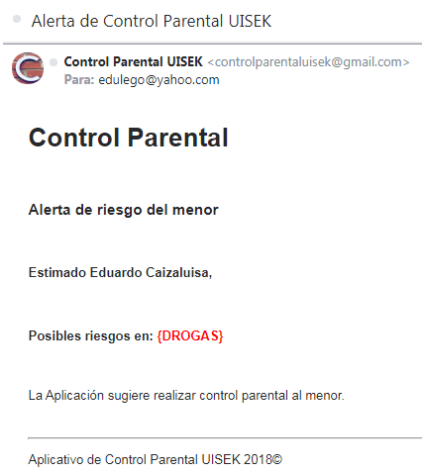

Figura 24. Captura del Correo de alerta lanzado por el aplicativo por DROGAS

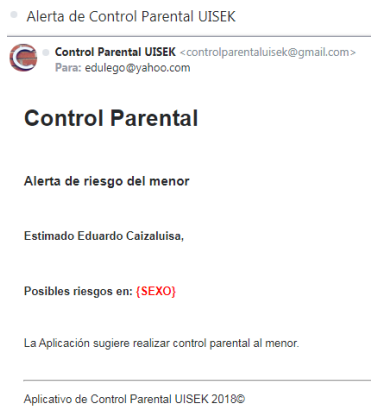

Figura 25. Alerta lanzado por el aplicativo al detectar mensajes del tipo SEXO

4.5 Resumen de resultados

Analizando los resultados obtenidos (tabla 23) mediante el proceso de validación de resultados en el 1er y 2do intento, se obtiene que el primer intento es el que entrega un mejor promedio en todas las categorías aplicadas y bajo las condiciones indicadas anteriormente, es decir utilizando el Algoritmo *Naíve Bayes*.

1er Intento de Validación		**2do Intento de Validación**	
BULLYING	91.66%	*BULLYING*	100%
DROGAS	3.70%	DROGAS	3.70%
SEXO	41.66%	SEXO	0%
NEUTRAL	100.0%	NEUTRAL	0%
PROMEDIO	**59.25%**		**25.92%**

Tabla 23. Resultado de clasificación de mensajes en porcentajes

4.6 Discusión de resultados

De los resultados obtenidos en la ejecución del primer intento de las pruebas para *bullying* 91.66%, *drogas* 3.70%, *sexo* 41.66% y *neutral* 100% podemos deducir que los pronósticos en conversaciones de los tipos **NEUTRAL** y ***BULLYING*** tienen un resultado alto, mientras que en la categoría **DROGAS** se tiene una predicción muy baja.

Por otro lado, en los resultados obtenidos para el segundo intento de validación, *bullying* obtiene una calificación perfecta, lo que no ocurre con las otras categorías, en

donde *sexo* y *neutral* obtienen 0% de certeza, no muy lejos *drogas* con el 3.70%, representando una certeza muy baja en relación con el primer intento de validación.

Hipotéticamente hablando, se puede dar algunas explicaciones a este fenómeno. Esto puede deberse a que *Twitter* cuenta con un pre filtrado para contenido inapropiado de acuerdo a su *Twitter-rules*[19], el cual evita conversaciones del tipo SEXO, VIOLENCIA, DROGAS, etc., además de que los mismos usuarios pueden reportar contenido inapropiado, discriminatorio, racista o que cause algún tipo de rechazo por su contenido a la comunidad. Otra explicación podría ser que en el proceso de colecta de datos (explicado en la sección *3.4.1 Datos Base de Entrenamiento*) se puede llegar a dar imprecisiones y ambigüedades en el contenido de los documentos clasificados en relación a la tabla 5 de palabras de filtros definidos, en este sentido, el algoritmo de clasificación retorna resultados distintos al esperado.

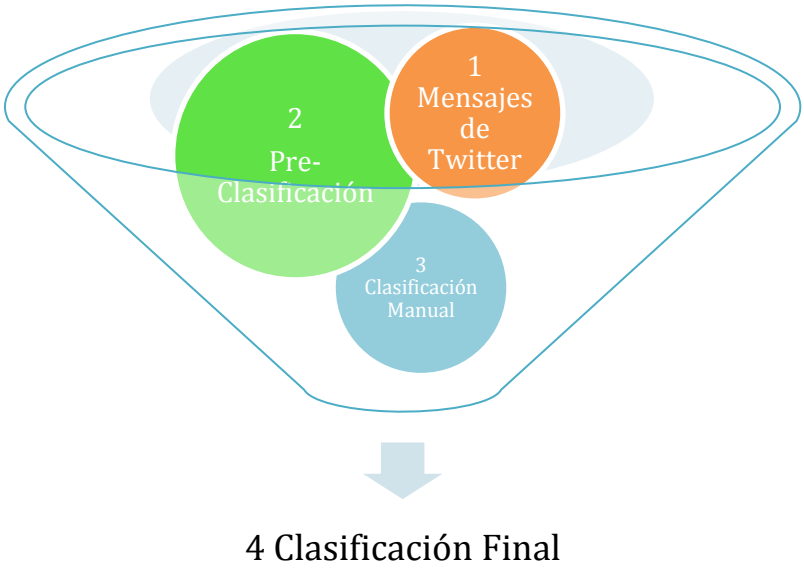

Figura 26. Diagrama ideal sugerido para la clasificación de mensaje

Con respecto a esta última hipótesis, para mejorar los resultados del pronóstico del aplicativo, se propone realizar una depuración manual (Hombre - Máquina) antes de ser catalogados definitivamente, como lo muestra el diagrama de la figura 26.

Este procedimiento lo recomienda Guyon et.al. (1996) para una minería de datos efectiva, ya que, en el proceso de clasificación de mensajes, pueden existir datos sucios o también ambigüedades producto de una mala recolección o producto también de la misma jerga sintáctica del idioma y sus modismos. Por ejemplo, no es lo mismo referirse a "eres un loco" pensando como una frase de bullying que refiriéndose a una frase de admiración sobre una actividad osada, todas estas inconsistencias pueden dar a lugar a que el programa obtenga falsos pronósticos.

[19] https://help.twitter.com/es/rules-and-policies/twitter-rules

Ahondando más, es necesario dar una explicación técnica de los resultados en los dos intentos de validación. Para esto, se utiliza el método *metrics.classification_report,* el cual muestra algunas estadísticas útiles en el tiempo de ejecución de la predicción. Este programa devuelve los resultados obtenidos en la tabla 24 y 25.

	Precisión	**Re-llamada**	**F1-score**
Bullying	0.75	0.95	0.84
Drogas	0.92	0.93	0.92
Neutral	0.97	0.95	0.96
Sexo	0.93	0.76	0.83
Promedio / Total	0.92	0.91	0.91

Tabla 24. Resultado *metrics.classification_report* en el primer intento de validación

	Precisión	**Re-llamada**	**F1-score**
Bullying	0.75	0.94	0.84
Drogas	0.98	0.97	0.97
Neutral	0.96	0.92	0.94
Sexo	0.91	0.74	0.82
Promedio / Total	0.90	0.89	0.89

Tabla 25. Resultado *metrics.classification_report* en el segundo intento de validación

Analizando el resultado obtenido en el primer intento de validación, aplicando el método *metrics.f1_score* en la conversación "hasta que tengas un orgasmo", el programa pronosticó la categoría NEUTRAL e indica que obtenemos un valor de 0.41 de precisión (*score*) tomando que cuenta que *F1_score*[20] alcanza su mejor valor cuando tiende a 1 y su peor valor cuando tiende a 0, aplicando una regla de tres para *f1_score*

[20] http://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html

tenemos que $F1\ Score = \frac{1}{0.41}\ \frac{100\%}{X}$ da un 41%. Tomando en cuenta que es un valor bajo, que no llega ni al 50% de certeza, con esto se puede suponer que el resultado no puede ser el mejor y podría descartarse con el fin de no lanzar falsas alertas.

En cambio cuando aplicamos el *metrics.f1_score* para el segundo intento de validación y la conversación "hasta que tengas un orgasmo" el programa pronosticó *BULLYING* y obtuvo un *score* de 0.89, haciendo una regla de tres simple obtenemos 89% de precisión, el cual se puede dar como un buen resultado, sin embargo, el pronóstico no es el adecuado.

Analizando los resultados de la tabla 23 después de la ejecución de los 2 intentos de validación, se puede visualizar que en cada intento se tuvo resultados diferentes y en cada una de las categorías el porcentaje obtenido varía. Es así que, en el primer y segundo intento para *BULLYING* se obtuvo un 91.66% y 100% respectivamente, mientras que para DROGAS se obtuvo un 3.70% para los dos intentos. En cambio, en las categorías SEXO y NEUTRAL se obtuvo un 41.66% y un 100% en el primer intento, mientras que in 0%, en ambas categorías, en el segundo intento, lo que indica que son resultados opuestos. Una hipótesis sobre esta diferencia, puede deberse a que se ejecutó sobre bancos de datos diferentes por la cantidad de documentos recolectados para cualquiera de los dos intentos ejecutados.

CAPÍTULO V

CONCLUSIONES Y TRABAJOS FUTUROS

5.1 Conclusiones y Recomendaciones

Conforme a los objetivos planteados, se culminó con el desarrollo de una aplicación para dispositivos móviles *Android* y la respectiva implementación en un dispositivo de prueba. Conforme a lo esperado, la aplicación móvil trabaja conjuntamente con el servidor remoto, el mismo que recibe los mensajes y realiza el procesamiento para clasificar el texto y obtener una predicción de los mensajes enviados, para de esta manera detectar posibles riesgos en los mensajes de *WhatsApp*. Durante las pruebas se verifica que el aplicativo realiza la respectiva notificación mediante un correo como un aporte y una ayuda para el control parental.

El aplicativo fue desarrollado e implementado bajo una arquitectura cliente/servidor respetando el diseño con el cual fue concebido para un despliegue en 4 capas: Servidor Remoto, Aplicativo *Web* y Base de Datos.

En base a la investigación sobre el Sistema Operativo *Android* y *WhatsApp*, en cuanto a la arquitectura y estructura, se pudo encontrar una estrategia de obtención de privilegios para obtener la base de datos de *WhatsApp*, ya que es un insumo importante de datos para el aplicativo.

Una vez que todos los componentes del aplicativo fueron desarrollados y correctamente implementados conforme a los objetivos propuestos, el aplicativo móvil en *Android* fue instalado exitosamente en un celular de pruebas y con esto se realizaron pruebas y evaluación de resultados.

Mediante el uso de la herramienta *Scikit-learn* y la selección del algoritmo *MultinomialNB* para la clasificación de los datos se realizó el entrenamiento del modelo para el pronóstico de un posible riesgo. Para esto, se entrenó el algoritmo utilizando mensajes de *Twitter*.

Con respecto a los resultados obtenidos en el primer intento de validación para los mensajes definidos para la ejecución de las pruebas, se tiene para *bullying* 91.66%, drogas 3.70%, sexo 41.66%, neutral 100%. Estos valores indican que se ve la necesidad de cambiar de estrategia de adquisición y clasificación de los datos para que éste llegue a tener una mejor asertividad en los resultados, la limpieza debe aplicar para todas las categorías.

Para que el aplicativo entregue un pronóstico libre de errores sobre un posible riesgo en el contenido de los mensajes, los datos que el algoritmo procesa deben estar

adecuadamente clasificados y en el mejor de los casos realizado una depuración de datos manual (Hombre – Máquina).

Durante el proceso del desarrollo del proyecto, se notó sobre la carencia de contar con un corpus relativo a palabras, frases, expresiones, etc. utilizadas en Latinoamérica y más específicamente en el Ecuador, mucho menos encontrar expresiones relacionadas a como se expresan los menores o jóvenes cuando se comunican entre ellos, dentro del contexto de las categorías definidas en este trabajo de investigación (SEXO, DROGAS, *BULLYING*, NEUTRAL), si bien es cierto la Real Academia de la Lengua cuenta con un corpus, este corpus no se adapta a el contexto de los mensajes intercambiados en el WhatsApp por los jóvenes, ni tampoco a las expresiones utilizadas en el Ecuador, razón por la cual se debió generar un propio corpus que pudo no ser el mejor.

Si bien es cierto, el aplicativo requiere tener un acceso *root* en el dispositivo móvil, es necesario advertir a los usuarios de los riesgos en cuanto a la brecha de seguridad que se abre al realizar este procedimiento. Por otro lado, puede existir una vulneración de la privacidad del menor de edad al analizar los datos del WhatsApp, desde el punto de vista jurídico, el Código de la Niñez y Adolescencia indica, en el Artículo 46 sobre las prohibiciones al *Derecho de la Información* de los menores cuando intercambian contenidos inadecuados para su desarrollo, así también el Artículo 78 habla del derecho a protección contra otras formas de abuso contra drogas, violencia o todo lo que pueda a dañar la integridad de los menores, en el Artículo 73 habla del deber de protección en el caso de maltrato ya que insta a toda persona a intervenir en el acto para proteger al menor en casos flagrantes de maltrato, abuso sexual, tráfico y explotación sexual, etc.

A manera de recomendación, cuando el aplicativo propuesto alcance un nivel de madurez tal como para entrar a producción, es necesario poner un *Acuerdo de Responsabilidad* que especifique sobre los riesgos, implicaciones de seguridad y privacidad, y potenciales peligros a los que se somete al usuario del dispositivo móvil al realizar el procedimiento de *root*, para que el usuario final esté consiente y de acuerdo con las instrucciones del *Acuerdo de Responsabilidad*.

5.2 Limitaciones

Para lograr obtener acceso a la base de datos del *WhatsApp* en la aplicación móvil, fue necesario por un lado realizar el *flasheo*[21] (Actualización o restablecimiento del *software* base en un dispositivo móvil con respecto al que vino de fábrica) y por otro, instalar un programa para la administración de accesos *root* (Obtención de los máximos privilegios en el dispositivo móvil) con la utilización de la aplicación *Kingo SuperUser* , por lo cual estamos permitidos de acceder al sistema de archivos restringido, en este caso la ruta que se requiere acceder se encuentra en: `/data/data/com.whatsapp/databases/msgstore.db`. Sin embargo, al realizar este paso, se está abriendo una brecha de seguridad en el dispositivo móvil, dejando vulnerable al dispositivo móvil, por lo cual terceros puede valerse para comprometer los datos del usuario.

El *corpus* generado para la solución y explicado en el apartado *3.4.1 Datos Base de Entrenamiento*, es meramente provisional para tratar de dar cumplimiento al objetivo del presente trabajo, en este sentido, es vital generar investigaciones en este tema, que doten de una *corpora* robusta y de calidad, relacionada a las expresiones habladas y escritas por la juventud actual en el Ecuador, y que sea la base de entrenamiento del modelo para la predicción sobre las categorías definidas (SEXO, DROGAS, *BULLYING*).

5.3 Trabajos futuros

El presente trabajo, servirá de base para implementar el monitoreo en otras redes sociales como *Facebook, Snapchat, Instagram*, etc. Para esto, se recomienda seguir procedimientos similares al estudio del presente trabajo, para de esta forma adquirir la base de datos, y enviarla a un servidor central para el procesamiento, análisis y entrega de resultados.

Por el momento el aplicativo está implementado para *Android*, sin embargo en trabajos futuros puede ser desarrollado e implementado para *iOS* en las diferentes versiones del sistema operativo y distintos modelos de dispositivos existentes en el mercado.

Con respecto a la arquitectura del presente desarrollo de software, se podría reutilizar la mayoría de los componentes, sin embargo, se puede comparar los resultados de otro tipo de algoritmo de clasificación como por ejemplo el algoritmo *Support Vector Machine* (*SVM*). Esto con el fin de si se pueden obtener mejores resultados de pronóstico, lo que ayudaría de gran manera para poder mejorar el aplicativo hasta llegar a tener un óptimo desempeño en la detección de amenazas y reducción de los riesgos contra el menor de edad.

[21] http://comodesbloquearcelular.com/flasheo-de-celulares/

Es necesario realizar un mecanismo de selección del corpus de las palabras mediante un estudio de campo con el fin de determinar el *corpus* de forma adecuada con respecto a la jerga con la que se comunican los jóvenes actualmente; para cada categoría SEXO, DROGAS, *BULLYING* y otras categorías que se definan para el entrenamiento del modelo.

Desarrollar un *corpus* relacionado al habla latinoamericano en especial con expresiones ecuatorianas, que sirva de insumo para otros trabajos de investigación relacionados a este tema para el trabajo de clasificación de textos y análisis lingüístico.

Para evitar abrir una brecha de seguridad en un dispositivo móvil con el procedimiento *root* planteado en el presente trabajo, es necesario realizar un estudio en donde se determine un procedimiento menos intrusivo con el cual se pueda adquirir los datos.

BIBLIOGRAFÍA

Abad-Grau, M. M., Ierache, J. S., & Cervino, C. (2007). Aplicación de redes bayesianas en el modelado de un sistema experto de triaje en servicios de urgencias médicas. In IX Workshop de Investigadores en Ciencias de la Computación.

Aguado Corman, A. (2016). Clasificación de actividades humanas en tiempo real a partir de representaciones en esqueleto.

Agustina, J. R. (2010). ¿Menores infractores o víctimas de pornografía infantil? Respuestas legales e hipótesis criminológicas ante el Sexting. Revista Electrónica de Ciencia Penal y Criminología.

Amato, G., Bolettieri, P., Costa, G., La Torre, F., & Martinelli, F. (2009, September). Detection of images with adult content for parental control on mobile devices?. In Proceedings of the 6th International Conference on Mobile Technology, Application & Systems (p. 35). ACM.

Arias Figueroa, M. A. (2016). Diseño e implementación de un algoritmo de clasificación de objetos peligrosos en radiografías utilizando el modelo bag of words.

Arias, F. G. (1999). El proyecto de investigación. Fidias G. Arias Odón.

Astray, G., Castillo, J. X., Ferreiro-Lage, J. A., Galvez, J. F., & Mejuto, J. C. (2010). Artificial neural networks: a promising tool to evaluate the authenticity of wine Redes neuronales: una herramienta prometedora para evaluar la autenticidad del vino. CyTA–Journal of Food, 8(1), 79-86

Banco Mundial (2018). Indicadores del desarrollo mundial. https://www.google.com.ec/publicdata/explore?ds=d5bncppjof8f9_&met_y=sp_pop_totl&hl=es&dl=es#!ctype=l&strail=false&bcs=d&nselm=h&met_y=sp_pop_totl&scale_y=lin&ind_y=false&rdim=country&idim=country:ECU&ifdim=country&hl=es&dl=es&ind=false. Acceso: 18-05-2018

Benalcázar Carrera, S. D., & Llumiquinga Lucero, G. V. (2016). Desarrollo de un sistema distribuido para la digitalización y procesamiento de cheques usando algoritmos de reconocimiento de dígitos manuscritos en la empresa Decisión CA (Bachelor's thesis, Universidad de las Fuerzas Armadas ESPE. Carrera de Ingeniería de Sistemas e Informática.).

Bosschert, B. (2017). Steal WhatsApp database (PoC). https://bas.bosschert.nl/steal-whatsapp-database/. Acceso: 11-12/2017

Çankaya, S., & Odabaşı, H. F. (2009). Parental controls on children's computer and Internet use. Procedia-Social and Behavioral Sciences, 1(1), 1105-1109.

Challenger-Pérez, I., Díaz-Ricardo, Y., & Becerra-García, R. A. (2014). El lenguaje de programación Python. Ciencias Holguín, 20(2), 1-12.

Christiane Fellbaum (1998, ed.) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.

Código de la Niñez y Adolescencia (2002), Asamblea Nacional del Ecuador Registro Oficial N° 737 del 03 enero 2003. Quito - Ecuador.

Código Orgánico Integral Penal. (2014), Asamblea Nacional del Ecuador Registro Oficial N° 180 del 10 febrero 2014. Quito - Ecuador.

Continente, X. G., Giménez, A. P., & Adell, M. N. (2010). Factores relacionados con el acoso escolar (bullying) en los adolescentes de Barcelona. Gaceta Sanitaria, 24(2), 103-108.

D'Adderio, D. (2017). More vs Cande: fragmentos de violencia en redes sociales. Letras.

Debnath, P., Haque, S., Bandyopadhyay, S., & Roy, S. (2016). Post-disaster Situational Analysis from WhatsApp Group Chats of Emergency Response Providers.

Domínguez, R. (2017). Aplicación de ciencia de datos para la creación de software predictivo de morbimortalidad materna en México.

Echeburúa, E., De Corral, P., & Amor, P. J. (2002). Evaluación del daño psicológico en las víctimas de delitos violentos. Psicothema, 14(Suplemento), 139-146.

Instituto Nacional de Estadísticas de España (2013), "Encuesta sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los Hogares (TIC-H)". http://www.ine.es/prensa/np803.pdf Acceso: 23/05/2018

Escolano Ruiz, F. (2003). Inteligencia artificial. Madrid: Thomson.

Fletcher, A. C., Steinberg, L., & Williams-Wheeler, M. (2004). Parental influences on adolescent problem behavior: Revisiting Stattin and Kerr. Child development, 75(3), 781-796.

García, C. G., Espada, J. P., G-Bustelo, C. P., & Lovelle, J. M. C. (2015). EL FUTURO DE APPLE: SWIFT VERSUS OBJECTIVE-C. Redes de Ingeniería.

García, J. (2016). Uso que los menores hacen de las redes sociales y control parental.

Garner (2017). Gartner Says Worldwide Sales of Smartphones Grew 9 Percent in First Quarter of 2017. https://www.gartner.com/newsroom/id/3725117. Acceso: 18-05-2018

González Sendino, R. (2017). Predicción de acciones de control aéreo (Doctoral dissertation, ETSI_Informatica).

Grant, K., & Haseman, C. (2013). Beginning Android programming: Develop and design. Pearson Education.

Graña Gil, J. (2002). Técnicas de análisis sintáctico robusto para la etiquetación del lenguaje natural. Procesamiento del lenguaje natural, nº 28 (mayo 2002); pp. 117-118.

Gupta, A. (2013). Java EE 7 essentials. O'Reilly Media.

Guyon, I. (2008). Introduction to Machine Learning. Slides and Videolecture.

Guyon, I., Matic, N., & Vapnik, V. (1996). Discovering Informative Patterns and Data Cleaning.

Hernández, E. A. (2009). Naíve Bayes Multinomial para Clasificación de Texto Usando un Esquema de Pesado por Clases.

Instituto Nacional de Estadística y Censos (2014). 1,2 millones de ecuatorianos tienen un teléfono inteligente (Smartphone). Quito - Ecuador: Instituto Nacional de Estadística y Censos. http://www.ecuadorencifras.gob.ec/12-millones-de-ecuatorianos-tienen-un-telefono-inteligente-smartphone/. Acceso: 12-11-2017

Introduction to modern information retrieval: G. Salton and M. McGill. McGraw-Hill, New York (1983)

Ivorra Méndez, E. (2014). Estudio sobre la protección de los padres hacia sus hijos en internet: control parental y otras estrategias, en Huelva capital, España (Doctoral dissertation, Universidad Internacional de Andalucía).

Jiménez, G. (2017). Te contamos de la nueva forma de extorsionar por WhatsApp. http://www.excelsior.com.mx/nacional/2017/02/27/1148946 Acceso: 25-10-2017

John Horton (2015) Android Programming for Beginners, Packt Publishing Ltd.

Kendall, K. E. (2005). Análisis y diseño de sistemas. Pearson educación.

Kuppusamy, K. S., Francis, L. M., & Aghila, G. (2013). Report: A Model for Remote Parental Control System Using Smartphones.

Larreátegui, G. B., & Sánchez, M. S. (2016). Implementación de una aplicación para control parental en dispositivos inteligentes. INVESTIGATIO RESEARCH REVIEW, (7), 101-115.

Rokach, L. & Maimon, O. (2008). Data mining with decision trees: theory and applications. World Scientific

López, I. P. V., Montalván, G. P. M., & Pin, M. V. S. SEXTING ENTRE ADULTOS JÓVENES DE LA UNIVERSIDAD TÉCNICA DE MANABÍ EN EL AÑO 2017 Autores e información del artículo.

Los Andes (2017). "Ballena azul", el macabro juego que circula por WhatsApp y pone en riesgo a los adolescentes: http://losandes.com.ar/article/-ballena-azul-el-macabro-juego-que-se-cuela-por-whatsapp-y-mata-a-chicos-y-adolescentes. Acceso: 16-10/2017

Lucio, P., & Vicente, J. (2018). The participation of the forensic physician in the crime scene. Medicina Legal de Costa Rica, 35(1), 102-114.

McCallum, A., & Nigam, K. (1998, July). A comparison of event models for naive bayes text classification. In AAAI-98 workshop on learning for text categorization (Vol. 752, No. 1, pp. 41-48).

Marcelo, J. F. y Martín, E. (2010). Protege a tus hijos de los riesgos de Internet y otras tecnologías. Madrid: Grupo Amaya, S. A.

MARÍA, J., & GINER, M. (2017). Visualización y seguimiento de acontecimientos en Twitter (Doctoral dissertation).

Mariño Solís, R. (2017) Predicción de la toxicidad y de la actividad antimicrobiana a partir de la secuencia aminoacídica.

Martínez, R. E. B., Ramírez, N. C., Mesa, H. G. A., Suárez, I. R., Trejo, M. D. C. G., León, P. P., & Morales, S. L. B. (2009). Árboles de decisión como herramienta en el diagnóstico médico. Revista médica de la Universidad Veracruzana, 9(2), 19-24.

Mendoza, M., Ortiz, I., & Rojas, V. (2011). Categorización de texto en bases documentales a partir de modelos computacionales livianos. Revista signos, 44(77), 251-274.

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (Eds.). (2013). Machine learning: An artificial intelligence approach. Springer Science & Business Media.

Miller, G.A. & Fellbaum, C. Lang Resources & Evaluation (2007)

Montoro, A. F. (2013). Python 3 al descubierto. RC Libros.

Montraveta, A. F., & Vázquez, G. (2010). La construcción del wordnet 3.0 en espanol. In La lexicografía en su dimensión teórica (pp. 201-220). Servicio de Publicaciones.

González, M. & Rodríguez, M. (1999). Lenguaje natural e indización automatizada.

La Nación (2017). San Juan: un chico intentó suicidarse siguiendo una consigna del mortal juego de "la ballena azul": http://www.lanacion.com.ar/2020242-san-juan-un-chico-intento-suicidarse-siguiendo-una-consigna-del-mortal-juego-de-la-ballena-azul. Acceso: 22-09-2017

Ofcom (2015). Ofcom report on internet safety measures. https://www.ofcom.org.uk/__data/assets/pdf_file/0020/31754/Fourth-internet-safety-report.pdf. Acceso: 27/12/2017

Ojeda-Pérez, J. E., Rincón-Rodríguez, F., Arias-Flórez, M. E., & Daza-Martínez, L. A. (2010). Delitos informáticos y entorno jurídico vigente en Colombia. Cuadernos de Contabilidad, 11(28).

Plaza Sacarrera, L. (2014). Análisis de polaridad en textos escritos en inglés y español (Master's thesis).

Policía Nacional del Ecuador (2017). Sujeto que extorsionaba mediante mensajes de WhatsApp fue detenido por la UNASE. http://www.policiaecuador.gob.ec/sujeto-que-extorsionaba-mediante-mensajes-de-whatsapp-fue-detenido-por-la-unase/. Acceso: 10-03-2017

Rodríguez, A. R. (2007) Extracción de Información Semántica a Partir de Categorías de Texto Estado del Arte.

Scikit-learn (2017). Machine Learning in Python. http://scikit-learn.org/stable/ Acceso: 07-01-2018

Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge university press.

Shutt-Aine, J., & Maddaleno, M. (2003). Salud sexual y desarrollo de adolescentes y jóvenes en las Américas: Implicaciones en programas y políticas. OPS.

Stattin, H., & Kerr, M. (2000). Parental monitoring: A reinterpretation. Child development, 71(4), 1072-1085.

Stuart, A.; Ord, K. (1994), Kendall's Advanced Theory of Statistics: Volume I— Distribution Theory

Teruel, M. (2015). Aprendizaje activo para clasificación de preguntas (Bachelor's thesis).

Unicef (2017). Dos tercios de los jóvenes en más de 18 países dicen haber sido víctimas de acoso escolar. https://www.unicef.es/prensa/dos-tercios-de-los-jovenes-en-mas-de-18-paises-dicen-haber-sido-victimas-de-acoso-escolar. Acceso: 26-02-2018

Valdés, J. T. (1987). Derecho informático (Vol. 102). Universidad nacional autónoma de México.

Vallejo, A. M. P. (2017). Bullying e Cyberbullying: Hoja de ruta y principales retos para la intervención. Pensar-Revista de Ciências Jurídicas, 22(1), 34-58.