



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



**TFG del Grado en Ingeniería
Informática**

título del TFG



Presentado por Nombre del alumno
en Universidad de Burgos — 27 de marzo
de 2019

Tutor: nombre tutor



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



D. nombre tutor, profesor del departamento de nombre departamento, área de nombre área.

Expone:

Que el alumno D. Nombre del alumno, con DNI dni, ha realizado el Trabajo final de Grado en Ingeniería Informática titulado título de TFG.

Y que dicho trabajo ha sido realizado por el alumno bajo la dirección del que suscribe, en virtud de lo cual se autoriza su presentación y defensa.

En Burgos, 27 de marzo de 2019

Vº. Bº. del Tutor:

Vº. Bº. del co-tutor:

D. nombre tutor

D. nombre co-tutor

Resumen

En este primer apartado se hace una **breve** presentación del tema que se aborda en el proyecto.

Descriptores

Palabras separadas por comas que identifiquen el contenido del proyecto Ej: servidor web, buscador de vuelos, android ...

Abstract

A **brief** presentation of the topic addressed in the project.

Keywords

keywords separated by commas.

Índice general

Índice general	III
Índice de figuras	IV
Índice de tablas	V
Introducción	1
Objetivos del proyecto	3
Conceptos teóricos	5
3.1. Secciones	5
3.2. Referencias	5
3.3. Imágenes	6
3.4. Listas de ítems	6
3.5. Tablas	7
Técnicas y herramientas	9
Aspectos relevantes del desarrollo del proyecto	11
Trabajos relacionados	13
Conclusiones y Líneas de trabajo futuras	19
Bibliografía	21

Índice de figuras

3.1. Autómata para una expresión vacía	6
--	---

Índice de tablas

3.1. Herramientas y tecnologías utilizadas en cada parte del proyecto	8
6.2. Objetivo de cada artículo	18

Introducción

Descripción del contenido del trabajo y del estructura de la memoria y del resto de materiales entregados.

Objetivos del proyecto

Este apartado explica de forma precisa y concisa cuales son los objetivos que se persiguen con la realización del proyecto. Se puede distinguir entre los objetivos marcados por los requisitos del software a construir y los objetivos de carácter técnico que plantea a la hora de llevar a la práctica el proyecto.

Conceptos teóricos

En aquellos proyectos que necesiten para su comprensión y desarrollo de unos conceptos teóricos de una determinada materia o de un determinado dominio de conocimiento, debe existir un apartado que sintetice dichos conceptos.

Algunos conceptos teóricos de \LaTeX ¹.

3.1. Secciones

Las secciones se incluyen con el comando `section`.

Subsecciones

Además de secciones tenemos subsecciones.

Subsubsecciones

Y subsecciones.

3.2. Referencias

Las referencias se incluyen en el texto usando `cite` [9]. Para citar webs, artículos o libros [3].

¹Créditos a los proyectos de Álvaro López Cantero: Configurador de Presupuestos y Roberto Izquierdo Amo: PLQuiz

3.3. Imágenes

Se pueden incluir imágenes con los comandos standard de \LaTeX , pero esta plantilla dispone de comandos propios como por ejemplo el siguiente:



Figura 3.1: Autómata para una expresión vacía

3.4. Listas de items

Existen tres posibilidades:

- primer item.
- segundo item.

1. primer item.
2. segundo item.

Primer item más información sobre el primer item.

Segundo item más información sobre el segundo item.

▪

3.5. Tablas

Igualmente se pueden usar los comandos específicos de \LaTeX o bien usar alguno de los comandos de la plantilla.

Herramientas	App	AngularJS	API REST	BD	Memoria
HTML5		X			
CSS3		X			
BOOTSTRAP		X			
JavaScript		X			
AngularJS		X			
Bower		X			
PHP			X		
Karma + Jasmine		X			
Slim framework			X		
Idiorm			X		
Composer			X		
JSON		X	X		
PhpStorm		X	X		
MySQL				X	
PhpMyAdmin				X	
Git + BitBucket		X	X	X	X
MikTeX					X
TeXMaker					X
Astah					X
Balsamiq Mockups		X			
VersionOne		X	X	X	X

Tabla 3.1: Herramientas y tecnologías utilizadas en cada parte del proyecto

Técnicas y herramientas

Esta parte de la memoria tiene como objetivo presentar las técnicas metodológicas y las herramientas de desarrollo que se han utilizado para llevar a cabo el proyecto. Si se han estudiado diferentes alternativas de metodologías, herramientas, bibliotecas se puede hacer un resumen de los aspectos más destacados de cada alternativa, incluyendo comparativas entre las distintas opciones y una justificación de las elecciones realizadas. No se pretende que este apartado se convierta en un capítulo de un libro dedicado a cada una de las alternativas, sino comentar los aspectos más destacados de cada opción, con un repaso somero a los fundamentos esenciales y referencias bibliográficas para que el lector pueda ampliar su conocimiento sobre el tema.

Aspectos relevantes del desarrollo del proyecto

Este apartado pretende recoger los aspectos más interesantes del desarrollo del proyecto, comentados por los autores del mismo. Debe incluir desde la exposición del ciclo de vida utilizado, hasta los detalles de mayor relevancia de las fases de análisis, diseño e implementación. Se busca que no sea una mera operación de copiar y pegar diagramas y extractos del código fuente, sino que realmente se justifiquen los caminos de solución que se han tomado, especialmente aquellos que no sean triviales. Puede ser el lugar más adecuado para documentar los aspectos más interesantes del diseño y de la implementación, con un mayor hincapié en aspectos tales como el tipo de arquitectura elegido, los índices de las tablas de la base de datos, normalización y desnormalización, distribución en ficheros³, reglas de negocio dentro de las bases de datos (EDVHV GH GDWRV DFWLYDV), aspectos de desarrollo relacionados con el WWW... Este apartado, debe convertirse en el resumen de la experiencia práctica del proyecto, y por sí mismo justifica que la memoria se convierta en un documento útil, fuente de referencia para los autores, los tutores y futuros alumnos.

Trabajos relacionados

En este apartado se comentará el estado del arte de la materia y algunos trabajos relacionados.

Para comenzar comentaremos que hay dos grupos de investigación importantes en este campo de estudio. El primer grupo sería el encabezado por el autor M. A. Little de la Universidad de Aston, Birmingham, Inglaterra. Este autor a su vez encabeza un proyecto para la obtención de una gran base de datos de 10000 audios obtenidos de pacientes con Parkinson. Esta iniciativa es conocida como *<Parkinson Vocie Initiative>*. Este conjunto de autores tiene varios artículos [7], [8] y el más citado [4]. Estos artículos aportan un buen análisis sobre las características a extraer de los audios (lineales y no lineales) y algoritmos de selección de características. Los aspectos más relevantes de cada artículo son los siguientes:

- *Suitability of dysphonia measurements for telemonitoring of Parkinson's disease* [4].

Este artículo trata sobre la aplicación de diferentes medidas estándares (lineales) y no estándares (no lineales) de disfonía para la clasificación automática de personas con Parkinson y personas sanas. Además añaden otra nueva medida calculada en este mismo paper llamada PPE, *pitch period entropy*. Se centra en responder a la pregunta de qué características son las mejores para la detección del Parkinson. Para ello utiliza únicamente audios de vocales sostenidas (28 PD y 8 HC). Se calculan un total de 17 medidas de disfonía de cada audio usando diferentes software como Praat **CITAR**. Posteriormente se hace un análisis de correlación entre las medidas obtenidas y de aquellos pares que tienen un coeficiente de correlación mayor del 95 % se

elimina una. Después de este análisis se obtienen 10 características: jitter (diferencia absoluta absoluta), jitter (media entre ciclos), APQ, shimmer (calculado como la diferencia absoluta promedio entre las amplitudes de los períodos consecutivos), HNR, NHR, RPDE, DFA, la dimensión de correlación y el comentado PPE.

En el siguiente paso del proceso utilizan el algoritmo de aprendizaje supervisado SVM con kernel de base radial para construir el modelo. Esto se realiza con cada uno de los 1023 diferentes subconjuntos posibles de las 10 características, $\sum_{i=1}^{10} \binom{10}{i}$, para encontrar el mejor subset posible. Esto es debido a que los autores consideran que es un número pequeño de conjuntos y se puede hacer búsqueda exhaustiva. Se llega a la conclusión que el **mejor subset de características es el formado por HNR, RPDE, DFA, y PPE** que devuelve una precisión del **91 %**, seguido en precisión por el subset completo de las 10 características.

- *Accurate Telemonitoring of Parkinsons Disease Progression by Noninvasive Speech Tests* [7].

En este artículo trata sobre cómo **el objetivo de la clasificación del Parkinson puede ser de regresión**. En vez de clasificar entre personas sanas y con PD, lo que se hace es medir el nivel de parkinson de una persona usando la escala UPDRS (*Unified Parkinson's Disease Rating Scale*). Para ello saca un total de 16 características de audios con vocales sostenidas, de las que se hace un análisis de correlación pero no se elimina ninguna.

Para la construcción de modelos se utilizan 4 técnicas diferentes: 3 de ellas de regresión lineal (LS, IRLS y LASSO) y una de regresión no lineal (CART's). Se llega a la conclusión de que los métodos lineales no dan malos resultados, siendo el IRLS el mejor de los 3. Sin embargo, el que mejor precisión da es el método CART. Los errores son de 8.47 ± 0.17 para UPDRS total con IRLS y de 7.52 ± 0.25 usando el método CART.

A parte de estos resultados, en el paper se realiza un análisis de la correlación de las características fijándose en los coeficientes devueltos por el método LS. En ellos se puede ver como las características altamente correlacionadas tienen magnitud similar y signo opuesto.

- *Novel Speech Signal Processing Algorithms for High-Accuracy Classification of Parkinson's Disease* [8].

En este paper el objetivo es clasificar entre personas sanas y con PD a partir de audios de vocales sostenidas **utilizando un gran conjunto**

de medidas de disfonía extraídas de los audios. Un punto novedoso de este experimento, es que hasta ese momento se habían elegido conjuntos de pocas (<20) medidas de disfonía y medido su correlación. En este artículo trata un total de 132 medidas lineares y no lineares de cada audio a las que se aplicará diferentes algoritmos de selección de características (algoritmos FS) para elegir varios subsets (uno por cada algoritmo). Los algoritmos de selección de características han sido: LASSO, RELIEF, mRMR y LLBFS. Cada algoritmo de selección de ccas ha elegido un subset diferente, siendo básicamente medidas como jitter y shimmer, variantes de medidas de ruido, MFCCs y medidas no lineales. Se ha analizado el número de características que comprende cada subset y se ha llegado a que cada subset tendrá 10 medidas únicamente. Según se explica, esto es debido a que usando más de 22 características se tiende al sobreajuste y que tampoco hay mucha mejoría usando 22 características en lugar de 10.

Posteriormente cada uno de esos subset de características ha sido usado para construir dos clasificadores con dos métodos distintos: random forest y SVM con kernel Gaussiano.

Se obtienen resultados que mejoran cualquier artículo de de accuracy del 97.7 % utilizando las 132 características con SVM. Sin embargo y como acabamos de explicar anteriormente, en el mismo paper acusa del accuracy tan alto a un posible sobreajuste al usar tantas características. Por lo explicado anteriormente, cada algoritmo FS escoge un subset de 10 medidas elegidas. El mejor resultado se da para el subconjunto de características escogido por el algoritmo RELIEF y usando SVM, presentando una accuracy del 98.6 % frente al RF que con el mismo subset llega a un accuracy del 93.5 %.

Para terminar con el resumen de este paper, indicar que la importancia está en que se trabaja con un número grande de medidas (132), llegando a la conclusión de que los subconjuntos idóneos tendrán alrededor de 10 características ya que tienen resultados parecidos a los subsets grandes pero generalizarán mejor que éstos.

El segundo grupo de investigación y para nosotros más importante, ya que seguimos los pasos para la extracción de características, es el liderado por J. R. Orozco-Arroyave de la Universidad de Antioquía, Colombia. Estos artículos son más recientes que los de el anterior grupo de investigación y, aunque tienen menos repercusión, son autores que trabajan y colaboran con el MIT (Massachusetts Institute of Technology) el cual tiene gran prestigio. El artículo más importante del grupo y que aporta una nueva visión a esta materia es [6]. Esta nueva visión es debido a que toma en cuenta más audios

a parte de únicamente de la vocal sostenida. También tiene en cuenta varios idiomas. Es precisamente de éste artículo del que hemos querido seguir los pasos a la hora de determinar qué características sacar de los audios y cómo sacarlas. También valorar y agradecer a este grupo de investigación la cesión de los audios descritos en [5], un corpus de audios de vocales sostenidas, monólogos, frases y palabras recopilados de un total de 100 personas (50 sanas y 50 PD) que contiene aproximadamente 4200 audios. Analizamos el paper más importante de este grupo:

- *Automatic detection of Parkinson's disease in running speech spoken in three different languages* [6].

En este artículo clasifica entre personas sanas y con PD y se incorporan a la materia varias novedades relacionadas con los tipos de audios a utilizar.

La primera de ellas es utilizar audios de monólogos, texto leído y palabras aparte únicamente de la vocal sostenida. Esto es debido a que la mayoría de los experimentos se basan en vocales sostenidas, obviando los consejos de los neurólogos respecto a que las consonantes requieren un mejor control de los movimientos de órganos (la lengua por ejemplo) y de ahí se puede sacar mucha más información. Por ello se utiliza un método para la caracterización de las señales de voz, basado en la segmentación automática de las expresiones en cuadros con voz y sin voz.

La otra gran novedad que incorpora es la utilización de corpus de audios de varios idiomas a la vez, en este caso 3 diferentes. Esto se utiliza, por ejemplo, para lo que ellos mismos llaman cross-language experiments. En estos experimentos, el entrenamiento del modelo se hace en un idioma y el testeo con otro idioma diferente.

En este paper se realizan diferentes experimentos, construyendo varios modelos para cada tipo de audio y así evaluar qué tipo de audio es el mejor para la clasificación. Se comienza haciendo un preprocesado de los audios donde los silencios iniciales y finales son eliminados. Se prosigue dividiendo los audios en partes con voz y partes sin voz (estos frames con voz y sin voz se corresponden a los frames donde el software Praat detecta que hay discurso y donde no lo hace). Para acabar el preprocesado se eliminan los fragmentos menores a 40 ms.

A continuación empezamos con la extracción de características de los audios. Se sacan 3 grupos de medidas diferentes con las que se construyen 3 modelos diferentes: 2 sacando diferentes tipos de medidas de los fragmentos con voz y 1 sacando medidas de los fragmentos sin voz. De los fragmentos con voz se sacan por un lado 64 medidas

prosódicas (jitter, shimmer...) con sus estadísticos (media, desviación, curtosis y oblicuidad). Por otro se sacan NHR, NNE, GNE, $F1$, $F2$ y las 12 MFCC con sus estadísticos. De los frames sin voz se hace un análisis cepstral y de energía para calcular las 12 MFCC y 25 BBE con sus estadísticos.

Después se hace la clasificación utilizando SVM con kernel Gaussiano que utiliza grid search para los parámetros C que mide la estrictez del margen y γ que es un parámetro del kernel Gaussiano utilizando como medida objetivo accuracy. Se utiliza una técnica de validación cruzada 10-fold.

Como resultado obtenemos 3 modelos diferentes para cada tipo de audio. De cada modelo saca 4 medidas para poderles comparar: *accuracy*, *sensitivity*, *specificity* y *AUC* (área bajo la curva de ROC). En el paper compara los experimentos por tipo de audio. El mejor rendimiento lo consiguen el texto leído ya que consigue, para el idioma castellano y con el modelo construido con las medidas extraídas de los frames sin voz, un accuracy del 97 % y un AUC del 99 %.

Ahora trataremos dos trabajos de una base de datos adquirida. Estos trabajos corresponden a Giovanni Dimauro de la Universidad de Bari, Italia. En el segundo de estos trabajos se aborda una perspectiva diferente para la diagnosis de Parkinson utilizando sistemas Speech-To-Text:

- *VoxTester, software for digital evaluation of speech changes in Parkinson disease* [2].

Este artículo describe la realización de una herramienta software para la ayuda en la evaluación de los cambios y variaciones en la voz de pacientes con la enfermedad del Parkinson. El funcionamiento de la herramienta es el siguiente: insertamos una serie de audios del paciente, la herramienta analiza el audio y nos devuelve 4 gráficas que un experto deberá interpretar para la evaluación de la enfermedad. Las gráficas de salida son las siguientes: gráfica del DDK rate, gráfica de la intensidad vocal y duración del discurso, gráfica del espectro de frecuencias vocal y gráfica del nivel de presión vocal. Éstos parámetros sacados de la voz son importantes ya que pueden ser indicadores de la enfermedad del Parkinson. Por ejemplo, un rango de frecuencias de la voz pequeño, una frecuencia fundamental baja o un decaimiento notable de la intensidad de la voz en el discurso serán una de las características del discurso para personas con esta enfermedad.

En este experimento no se utiliza en ningún momento técnicas de clasificación ni tampoco la herramienta discierne por ella misma si la

Artículos	Clasificación PD/HC	Regresión Mapeo UPDRS	Otro análisis
[4] [8] [6]	X		
[7]		X	
[2] [1]			X

Tabla 6.2: Objetivo de cada artículo

persona del audio tiene una enfermedad o en qué grado. Lo único que hace es analizar la onda sonora y mostrar algunas características, por lo que no se alinea del todo con nuestros objetivos.

- *Assessment of Speech Intelligibility in Parkinson's Disease Using a Speech-To-Text System* [1].

Este experimento trata de resolver el problema del anterior (valuación de los cambios y variaciones en la voz de pacientes con la enfermedad del Parkinson) con otra línea. Para ello utilizará sistemas de reconocimiento de voz y transcripción a texto (Speech-To-Text Systems) para evaluar la inteligibilidad del discurso de las personas con Parkinson. El proceso será el siguiente: la persona con Parkinson habla a un sistema STT, el texto de la transcripción se pasa a un programa que evalúa su acierto y finalmente ese programa devuelve el porcentaje de fallo en el reconocimiento de voz respecto a la frase objetivo. El objetivo de este proyecto ha sido comparar los fallos de reconocimiento en las palabras entre 3 grupos de personas: jóvenes sanos, mayores sanos y pacientes de la enfermedad del Parkinson. Los resultados obtenidos es que había mucho mas fallo en el reconocimiento de palabras en los pacientes de la enfermedad del Parkinson.

Conclusiones y Líneas de trabajo futuras

Todo proyecto debe incluir las conclusiones que se derivan de su desarrollo. Éstas pueden ser de diferente índole, dependiendo de la tipología del proyecto, pero normalmente van a estar presentes un conjunto de conclusiones relacionadas con los resultados del proyecto y un conjunto de conclusiones técnicas. Además, resulta muy útil realizar un informe crítico indicando cómo se puede mejorar el proyecto, o cómo se puede continuar trabajando en la línea del proyecto realizado.

Bibliografía

- [1] Giovanni Dimauro, Vincenzo Di Nicola, Vitoantonio Bevilacqua, Danilo Caivano, and Francesco Girardi. Assessment of speech intelligibility in parkinson's disease using a speech-to-text system. *IEEE Access*, pages 22199–22208, 2017.
- [2] Giovanni Dimauro, Vincenzo Di Nicola, Vitoantonio Bevilacqua, Francesco Girardi, and Vito Napoletano. Voxtester, software for digital evaluation of speech changes in parkinson disease. *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, pages 1–6, 2016.
- [3] John R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, 1992.
- [4] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig. Suitability of dysphonia measurements for telemonitoring of parkinson's disease. *IEEE Transactions on Biomedical Engineering*, 56:1015–1022, Enero 2009.
- [5] J. R. Orozco-Arroyave, J. D. Arias-Londoño, J. F. Vargas-Bonilla, M. González-Rátiva, and E. Nöth. New spanish speech corpus database for the analysis of people suffering from parkinson's disease. *Proceedings of the 9th Language Resources and Evaluation Conference (LREC)*, pages 342–347, 2014.
- [6] J. R. Orozco-Arroyave, F. Hönl, J. D. Arias-Londoño, J. F. Vargas-Bonilla, K. Daqrouq, S. Skodda, J. Rusz, and E. Nöth. Automatic detection of parkinson's disease in running speech spoken in three different languages. *The Journal of the Acoustical Society of America*, 139:481–500, Enero 2016.

- [7] A. Tsanas, M. A. Little, P. E. Mcsharry, and L. O. Ramig. Accurate telemonitoring of parkinsons disease progression by noninvasive speech tests. *IEEE Transactions on Biomedical Engineering*, 57:884–893, Abril 2010.
- [8] A. Tsanas, M. A. Little, P. E. Mcsharry, J. Spielman, and L. O. Ramig. Novel speech signal processing algorithms for high-accuracy classification of parkinson’s disease. *IEEE Transactions on Biomedical Engineering*, 59:1264–1271, Mayo 2012.
- [9] Wikipedia. Latex — wikipedia, la enciclopedia libre, 2015. [Internet; descargado 30-septiembre-2015].