

## TABLA DE CONTENIDOS

	página
<b>Dedicatoria</b>	<b>I</b>
<b>Agradecimientos</b>	<b>II</b>
<b>Tabla de Contenidos</b>	<b>III</b>
<b>Índice de Figuras</b>	<b>VI</b>
<b>Índice de Tablas</b>	<b>VII</b>
<b>Resumen</b>	<b>VIII</b>
<b>Resumen</b>	<b>IX</b>
<b>1. Introducción</b>	<b>10</b>
1.1. Descripción del Contexto . . . . .	11
1.2. Descripción del Problema . . . . .	11
1.3. Objetivos . . . . .	11
1.3.1. Objetivo General . . . . .	11
1.3.2. Objetivos Específicos . . . . .	11
1.4. Alcances del Proyecto . . . . .	12
1.5. Estructura del Documento . . . . .	12
<b>2. Antecedentes</b>	<b>13</b>
2.1. Reconocimiento de Voz . . . . .	13
2.1.1. La Voz . . . . .	13
2.1.2. Proceso de Digitalización de la Voz . . . . .	14
2.1.3. Técnicas de Extracción de Características . . . . .	15
2.2. Redes Neuronales Artificiales . . . . .	18
2.2.1. Origen . . . . .	18
2.2.2. Tipos de Redes Neuronales . . . . .	22
2.2.3. Redes Neuronales para Reconocimiento de Voz . . . . .	24
2.3. Búsqueda en Espacios Métricos . . . . .	26

2.3.1.	Levenshtein . . . . .	27
2.3.2.	Índices Basados en Pivotes . . . . .	28
2.3.3.	Búsqueda de los Cercanos . . . . .	28
2.3.4.	Selección Espacial de Pivotes Dispersos . . . . .	30
<b>3.</b>	<b>Análisis del Problema</b>	<b>31</b>
3.1.	Trabajos Existentes . . . . .	31
3.2.	Propuesta de Solución . . . . .	32
<b>4.</b>	<b>Desarrollo</b>	<b>33</b>
4.1.	Arquitectura de Software . . . . .	33
4.2.	Etapas del Proyecto . . . . .	35
4.3.	Metodología de Desarrollo . . . . .	35
4.3.1.	Entorno de Trabajo . . . . .	36
4.4.	Modelo de la Red Neuronal . . . . .	37
4.4.1.	Entradas . . . . .	37
4.4.2.	Salidas . . . . .	38
4.4.3.	Capas Ocultas . . . . .	38
4.4.4.	Capa de Clasificación Temporal . . . . .	39
4.5.	Conjunto de Datos de Entrenamiento . . . . .	41
4.6.	Entrenamiento . . . . .	43
4.7.	Corrector de Palabras . . . . .	44
4.8.	Implementación de la API . . . . .	46
<b>5.</b>	<b>Pruebas</b>	<b>48</b>
5.1.	Pruebas de Entrenamiento . . . . .	48
5.2.	Pruebas de Reconocimiento . . . . .	50
5.3.	Pruebas de Corrección de Palabras . . . . .	52
5.4.	Resultados . . . . .	55
<b>6.</b>	<b>Conclusiones y Trabajo futuro</b>	<b>57</b>
6.1.	Conclusión . . . . .	57
6.2.	Trabajo Futuro . . . . .	58
	<b>Glosario</b>	<b>59</b>

**Anexos**

<b>A: Formato de Documentos Oficiales Utilizados</b>	<b>64</b>
A.1. Formato de Listado de Requisitos . . . . .	65
A.2. Formato Consentimiento Informado . . . . .	68
A.3. Firma Aprobación de Requisitos . . . . .	69
A.4. Documentación API . . . . .	72
<b>B: Historias de Usuario</b>	<b>76</b>
<b>C: Pruebas de Corrección</b>	<b>77</b>

## ÍNDICE DE FIGURAS

	página
2.1. Diagrama en Bloques para la obtención de los MFCC. . . . .	16
2.2. Banco de Filtros en Escala Mel. . . . .	18
2.3. Neurona. . . . .	19
2.4. Representación del Perceptrón de Rosenblatt. . . . .	20
2.5. Modelo Estructural de una Red Neuronal Multicapas. . . . .	21
2.6. Celda de Memoria LSTM. . . . .	25
2.7. Red LSTM. . . . .	26
3.1. Etapas de Reconocimiento de Voz. . . . .	32
4.1. Arquitectura de la API. . . . .	34
4.2. Estructura Base de Datos. . . . .	42
5.1. Porcentaje de error vs correcciones en el dataset de prueba. . . . .	49
5.2. Error en Entrenamiento. . . . .	50
5.3. Prueba de Reconocimiento. . . . .	51
5.4. Costo de Corrección. . . . .	53
5.5. Histograma del largo de palabras del diccionario utilizado. . . . .	54

## ÍNDICE DE TABLAS

	página
2.1. Características de la Voz. . . . .	14
2.2. Comparativa de los tipos de redes neuronales. . . . .	23
4.1. Etapas del Proyecto. . . . .	35
4.2. Estructura Archivo Hablantes. . . . .	42
4.3. Sonido de letras críticas en Chile. . . . .	43
4.4. Extracto Matriz de penalización. . . . .	45
5.1. Iteraciones vs correcciones. . . . .	49
5.2. Tiempo de reconocimiento. . . . .	51
5.3. Extracto del Conjunto de Palabras de Prueba. . . . .	52
5.4. Conjunto de Palabras de Prueba. . . . .	54