

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO INGENIERIA COMERCIAL
SANTIAGO - CHILE



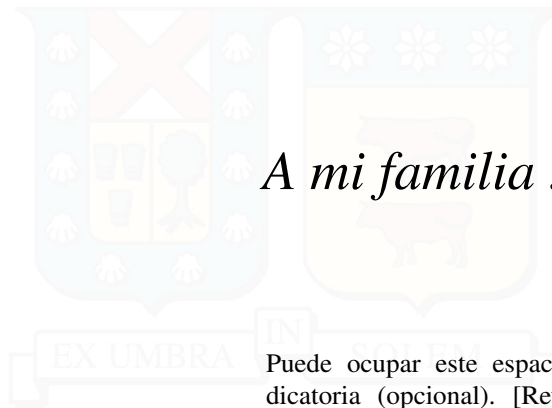
**TÍTULO DE MEMORIA (EL TÍTULO SÓLO PUEDE TENER UN MÁXIMO DE 3
LÍNEAS)**

SANTIAGO JESÚS VASCONCELLO ACUÑA

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO COMERCIAL

PROFESOR GUÍA : SR. PABLO ISLA
PROFESOR CORREFERENTE : SR. THIERRY DE SAINT PIERRE.

Diciembre 2023



A mi familia . . .

Puede ocupar este espacio para escribir una dedicatoria (opcional). [Revise el archivo maestro `memoria.tex` para modificar / eliminar esta sección.]

(AGRADECIMIENTOS) [Título es opcional]

Agradezco a quienes contribuyeron para ir mejorando esta plantilla hecha en L^AT_EX.

Los aportes y comentarios de distintas personas en el [Departamento de Industrias](#) fueron muy útiles para que este documento puede ser ocupado para mejorar la presentación de tesis y memorias del Departamento (y la universidad).

Para el impaciente ...

Por favor ocupar `git` :

```
git clone https://github.com/jaimercz/utfsm-thesis.git
```

Para los interesados en `git` revisar [1].

Abra el archivo de configuración `config.tex` para cambiar título, autor, fecha, etc. de la portada y del documento en general.

Abra y compile el documento maestro `memoria.tex` .

```
$ pdflatex memoria.tex
$ biber memoria
$ pdflatex memoria.tex
$ pdflatex memoria.tex
```

Esta version ocupa `biber` en lugar de `natbib` / `bibtex` :

Si hay errores, verifique primero que todos los paquetes L^AT_EX han sido instalados.

Si desea omitir alguna sección (dedicatoria, agradecimientos, etc.), revise el documento maestro `memoria.tex` y agregue o comente (o elimine) las líneas correspondientes.

Por ejemplo, para eliminar esta sección, borre las líneas:

```
_____ memoria.tex (extracto) _____
\section*{Agradecimientos}
\insertFile[plain]{agradecimientos}}
```

RESUMEN EJECUTIVO

Plantilla \LaTeX para las Memorias y Tesis del Departamento de Ingenieria Comercial, UTFSM.

Se incluyen también algunos ejemplos de cómo incorporar tablas y gráficos en distintas presentaciones respetando las Normas de Biblioteca para Memorias y Tesis de la UTFSM.

Palabras Clave. \LaTeX , Plantilla para Memoria, Departamento de Ingenieria Comercial, UTFSM.

¡Importante! [LEAME]

Impresión por un solo lado.

A partir del año 2016, el Departamento de Ingenieria Comercial sólo requiere la entrega digital de los archivos de memorias y tesis. Por este motivo, este documento está preparado para ser impreso por un solo lado de una hoja (“*oneside*”), y facilitar así su lectura en pantallas. Esta configuración es parte de archivo de clase `thesis_utfsm.cls`.

Codificación de caracteres.

Todos los archivos `*.tex` de esta plantilla han sido preparados ocupando la codificación de caracteres por defecto *unicode* (UTF-8). Windows (y algunas versiones de OSX) ocupan otro tipo de codificación (ej. *Windows-1252* o *Mac Roman*).

Si deseas ocupar esta plantilla y encuentras problemas con los caracteres acentuados, entonces puedes optar por una de estas tres alternativas:

- i) cambiar tu editor (TexMaker, TexStudio, TexShop, etc.) para que ocupe UTF-8 como codificación de caracteres por defecto; o
- ii) cambiar la codificación de cada documento `*.tex` para que ocupe la codificación nativa de tu sistema operativo; y, modificar el archivo `config.tex` la línea que dice:

OSX, Linux: `\usepackage[utf8x]{inputenc}`

Windows: `\usepackage[latin1]{inputenc}`

Overleaf: `\usepackage[utf8]{inputenc}` <https://overleaf.com>

- iii) escribir todo ocupando caracteres pre-acentuados (ej. `\'a` en lugar de á).

Recuerde:

Mezclar documentos de distintas codificaciones puede generar muchos problemas al momento de compilar.

ABSTRACT

This is a \LaTeX thesis template for the Departamento de Ingenieria Comercial, UTFSM. A few examples about the inclusion of figures and tables are also provided.

(The abstract can be edited by opening the file `includes/abstract.tex` .)

Keywords. \LaTeX , Thesis Template, Departamento de Ingenieria Comercial, UTFSM

Instrucciones para la Plantilla.

Editar el archivo `/includes/abstract.tex` para modificar los contenidos de esta sección.

Si no desea incluir un abstract, editar el archivo `/memoria.tex` , y comentar o borrar la sección que se muestra a continuación.

```
_____ /memoria.tex (extracto) _____  
\section*{ABSTRACT}  
\insertFile[plain]{abstract} % Archivo abstract.tex
```

Índice de Contenidos

1. Objetivo	1
1.1. Objetivo General	1
1.2. Objetivo Específico	1
1.3. Metodología	1
2. Estructura del proyecto	2

Índice de Tablas



Índice de Figuras

2.1. Estructura basica de la aplicación	2
---------------------------------------------------	---



1 | Objetivo

1.1. Objetivo General

Determinar los factores de riesgo que pueden llegar a influir, tanto de la creación como del uso de aplicaciones que utilicen modelos grandes de lenguaje (LLM) aplicados a la industria, usando de base el proyecto de búsqueda de jurisprudencia de los tribunales ambientales.

1.2. Objetivo Específico

1. Determinar una posible estructura de una aplicación usando LLM
2. Desarrollar los estados del Arte del uso de LLM y de los modelos generativo en si
3. Desarrollar los problemas que conlleva el uso de información para alimentar dichos problemas
4. Analizar un proceso de ETL de principio a fin para observar sus posibles riesgos

1.3. Metodologia

2 | Estructura del proyecto

Los Tribunales Ambientales son órganos jurisdiccionales especiales, sujetos a la superintendencia directiva, correccional y económica de la Corte Suprema, cuya función es resolver las controversias medioambientales de su competencia y ocuparse de los demás asuntos que la ley somete a su conocimiento *Ley Chile - Ley 20600 - Biblioteca del Congreso Nacional* [3]. Estos tribunales generan una cantidad de jurisprudencia que puede ser encontrada en su portal de consulta llamado buscador ambiental *Buscador Ambiental* [2].

El proyecto consiste en la generación de un chatbot en donde se pueda preguntar sobre la jurisprudencia de estos tribunales, aunque por razones de capacidad el chatbot se vea acotado solamente a las reclamaciones recibidas por el tribunal.

Por lo que, este proyecto consiste en un proceso de extracción de datos desde el buscador ambiental, transformación de estos datos para su utilización, generación de vectores de estos datos para que puedan interaccionar con la aplicación, carga de estos en una base de datos, para que después la aplicación pueda interactuar con ellos y mandando esa información a el LLM, siendo en este caso gpt-4 perteneciente a OpenAI.

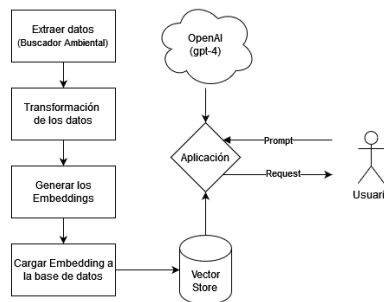
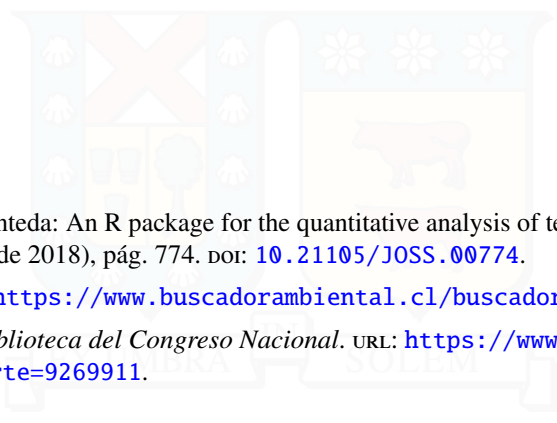


Figura 2.1: Estructura básica de la aplicación

(Fuente: Elaboración propia)

A partir de esta estructura mientras se avance en el desarrollo, se explicará parte por parte el proceso y con ello los riesgos de cada uno de ellos.

Bibliografía

- 
- [1] Kenneth Benoit y col. “quanteda: An R package for the quantitative analysis of textual data”. En: *Journal of Open Source Software* 3 (30 oct. de 2018), pág. 774. doi: [10.21105/JOSS.00774](https://doi.org/10.21105/joss.00774).
- [2] *Buscador Ambiental*. URL: <https://www.buscadorambiental.cl/buscador/#/>.
- [3] *Ley Chile - Ley 20600 - Biblioteca del Congreso Nacional*. URL: <https://www.bcn.cl/leychile/navegar?idNorma=1041361&idParte=9269911>.