

INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO



**MODELADO Y GENERACIÓN  
MUSICAL: UNA APROXIMACIÓN  
MEDIANTE REDES NEURONALES Y  
ENTRADAS MIDI**

TESINA

QUE PARA OBTENER EL TÍTULO DE

INGENIERÍA EN MECATRÓNICA

PRESENTA

SANTIAGO PÉREZ GUTIÉRREZ

ASESOR

WILMER EFRÉN PEREIRA GONZÁLEZ

# Tabla de contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Contexto . . . . .	1
1.2. Identificación del problema . . . . .	2
1.3. Objetivo . . . . .	2
1.4. Alcance . . . . .	2
1.5. Metodología . . . . .	4
1.5.1. Detectar brecha: . . . . .	5
1.5.2. Definir el problema: . . . . .	5
1.5.3. Explorar alternativas: . . . . .	5
1.5.4. Seleccionar plan: . . . . .	6
1.5.5. Documentación y Diseminación: . . . . .	7
1.6. Organización del documento . . . . .	7
<b>2. Análisis</b>	<b>9</b>
2.1. Requerimientos Funcionales . . . . .	9
2.2. Requerimientos No Funcionales . . . . .	10
2.3. Restricciones Funcionales . . . . .	11
2.4. Restricciones de Datos: . . . . .	11
2.5. Trabajos relacionados . . . . .	12
2.5.1. Revisión de la Literatura: . . . . .	12

<i>TABLA DE CONTENIDO</i>	2
2.5.2. Comparación de Modelos: . . . . .	13
<b>3. Diseño</b>	<b>15</b>
3.1. Arquitectura . . . . .	15
3.2. Soluciones Alternativas . . . . .	15
<b>4. Implementación</b>	<b>17</b>
4.1. Modelo óptimo . . . . .	17
<b>5. Pruebas y resultados</b>	<b>18</b>
5.1. Pruebas realizadas y sus resultados . . . . .	18
<b>6. Conclusiones</b>	<b>20</b>
6.1. Conclusiones pruebas de velocidad y posición . . . . .	20
<b>Referencias</b>	<b>22</b>

# Capítulo 1

## Introducción

### 1.1. Contexto

La música ha sido un reflejo de la evolución cultural y tecnológica de la humanidad desde tiempos inmemoriales. Desde las melodías simples creadas con instrumentos rudimentarios en culturas antiguas hasta las complejas composiciones orquestales del Renacimiento y más allá, la música ha servido como medio de expresión, comunicación y unión para las sociedades de todo el mundo.

Con la Revolución Industrial y los avances tecnológicos del siglo XX, la música experimentó una metamorfosis significativa. Surgieron nuevos instrumentos, métodos de grabación y formas de distribución, transformando radicalmente la industria musical. Una de estas revoluciones tecnológicas fue la introducción del formato MIDI en la década de 1980. Esta interfaz digital permitió una representación precisa de la música en términos de notas, duraciones, intensidades y otros parámetros. El MIDI no solo democratizó la producción musical, permitiendo a músicos sin acceso a grandes estudios producir música de alta calidad, sino que también sentó las

bases para la experimentación y manipulación digital de la música.

## **1.2. Identificación del problema**

En el siglo XXI, con la creciente omnipresencia de la inteligencia artificial en diversos campos, la música no ha sido una excepción. Las redes neuronales, en particular, han mostrado promesa en tareas que van desde la clasificación de géneros musicales hasta la generación automática de composiciones. Sin embargo, la combinación de estas técnicas avanzadas de IA con entradas MIDI para la generación y modelado de música es un área emergente y de rápido crecimiento, que busca cerrar la brecha entre la creatividad humana y la capacidad computacional.

## **1.3. Objetivo**

El propósito principal de este estudio es diseñar, implementar y evaluar una red neuronal avanzada capaz de procesar y interpretar fragmentos musicales presentados en formato MIDI. Más allá de simplemente generar una continuación basada en patrones aprendidos, el objetivo es que esta red neuronal produzca extensiones musicales que sean coherentes en estructura, melódicamente satisfactorias y estilísticamente alineadas con el fragmento inicial.

## **1.4. Alcance**

Este estudio abarcará varias áreas interdisciplinarias, incluyendo teoría musical, procesamiento digital de señales y técnicas avanzadas de aprendizaje automático. Los principales aspectos cubiertos incluyen:

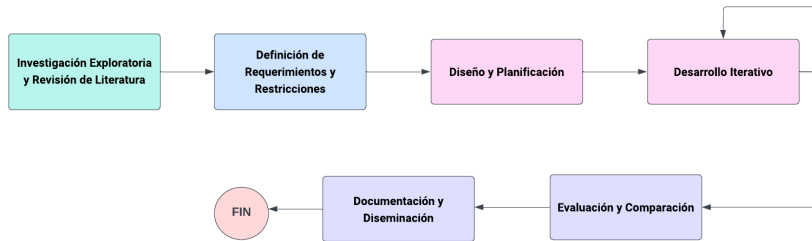
- **Investigación Teórica:** Se llevará a cabo una revisión exhaustiva de la literatura sobre generación musical asistida por IA, con un énfasis en trabajos anteriores que utilizan formatos MIDI y redes neuronales. Esta revisión ayudará a contextualizar el trabajo actual dentro del panorama más amplio del campo y a identificar áreas de oportunidad.
- **Desarrollo Técnico:** Se diseñará y construirá una red neuronal específicamente adaptada para procesar y generar música basada en entradas MIDI. Esto incluirá la selección de arquitecturas apropiadas, técnicas de preprocesamiento de datos y métodos de entrenamiento.
- **Comparación con Modelos Existentes:** Se comparará la eficacia del modelo desarrollado con modelos preentrenados existentes en términos de calidad musical y coherencia. Esto proporcionará un punto de referencia para evaluar las contribuciones y avances del modelo propuesto.
- **Evaluación:** Una vez desarrollada la red neuronal, se realizarán pruebas cuantitativas y cualitativas para evaluar la calidad de la música generada. Esto podría incluir métricas objetivas, como la coherencia melódica o armónica, así como evaluaciones subjetivas realizadas por expertos en música.
- **Aplicaciones Potenciales:** Se explorarán y discutirán posibles aplicaciones de la red neuronal desarrollada, desde herramientas de composición asistida para músicos hasta aplicaciones en educación musical o entretenimiento.
- **Limitaciones:** Aunque el estudio se esforzará por abordar una amplia variedad de géneros y estilos musicales, se reconocerá que no es posible cubrir todos los géneros y contextos. Además, se discutirán las

limitaciones inherentes a las redes neuronales y el formato MIDI en términos de representación y generación musical.

## 1.5. Metodología

Para abordar el problema se seguirán los siguientes pasos:

- Investigación Exploratoria y Revisión de Literatura: Revisar literatura académica y artículos de la industria. Estudiar modelos existentes y su funcionamiento. Identificar brechas y oportunidades para la innovación.
- Definición de Requerimientos y Restricciones: Definir los requerimientos funcionales y no funcionales. Identificar las restricciones técnicas y de datos.
- Diseño y Planificación: Seleccionar las arquitecturas de red neuronal y técnicas de procesamiento de datos MIDI adecuadas. Planificar las fases de desarrollo, pruebas y evaluación.
- Desarrollo Iterativo: Realizar pruebas cuantitativas y cualitativas. Comparar con modelos preentrenados y estudios relacionados. Recopilar feedback de usuarios y expertos.
- Documentación y Disseminación: Redactar informes detallados y documentos de investigación. Preparar presentaciones para compartir los hallazgos.



**Figura 1.4.1:** Metodología de Trabajo

### 1.5.1. Detectar brecha:

Se realizará una revisión sistemática de la literatura para identificar las lagunas en la investigación actual relacionadas con la generación de música utilizando IA y MIDI. Esta revisión no solo se centrará en lo que se ha hecho, sino también en lo que falta en el campo, tales como limitaciones en estilos musicales, diversidad de datos y rendimiento de los modelos existentes.

### 1.5.2. Definir el problema:

Con base en la revisión de la literatura, se establecerá una definición clara del problema que incluirá los desafíos específicos a abordar con la investigación. Esto incluirá la determinación de los requisitos técnicos y musicales para el modelo de red neuronal y la definición de qué constituye una generación musical exitosa

### 1.5.3. Explorar alternativas:

Investigar una variedad de arquitecturas de red neuronal, como las redes neuronales convolucionales (CNN), redes recurrentes (RNN), y las redes generativas adversarias (GAN), para determinar cuál es más adecuada para el modelado de secuencias musicales MIDI. Además, se evaluarán diferentes



métodos de preprocesamiento de datos MIDI y se determinará cómo influirán en la calidad de la salida musical.

#### **1.5.4. Seleccionar plan:**

Seleccionar la arquitectura de red neuronal más prometedora y desarrollar un plan para su implementación. Esto incluirá:

- Selección de Datos: Identificar y curar un conjunto de datos de entrenamiento de alta calidad que sea representativo de los estilos musicales que el modelo pretende generar.
- Desarrollo y Entrenamiento: Implementar la red neuronal seleccionada y entrenarla con el conjunto de datos. Esto también incluirá la optimización de hiperparámetros y la evaluación de la red durante el entrenamiento para asegurar la convergencia.
- Validación y Pruebas: Validar el modelo contra un conjunto de datos de prueba que no se haya utilizado durante el entrenamiento. Esto incluirá pruebas de robustez y la capacidad del modelo para generalizar a partir de nuevas entradas MIDI.
- Comparación de Modelos: Ejecutar una comparación detallada con modelos preentrenados existentes para evaluar el rendimiento relativo. Se utilizarán métricas como la precisión melódica, armónica y rítmica, así como la originalidad y la calidad estética percibida.
- Evaluación del Usuario: Realizar evaluaciones cualitativas con músicos y compositores para obtener retroalimentación sobre la utilidad y la musicalidad de las composiciones generadas.

### 1.5.5. Documentación y Diseminación:

Documentar todos los procesos y resultados de manera detallada para permitir la reproducibilidad y proporcionar una base sólida para futuras investigaciones.

## 1.6. Organización del documento

El documento está estructurado de la siguiente forma:

- **Introducción:** Descripción de la situación, contexto, causas y metodología a seguir para llegar a la solución y concluir con el objetivo de la tesina.
- **Análisis del problema:** Características generales que debe cumplir la tesina, así como las restricciones que presenta para el desarrollo, algunos trabajos que realizan una función similar y las principales diferencias que presenta la tesina con respecto a los trabajos relacionados.
- **Definición del sistema a implementar:** Detalles de la arquitectura bajo la cual funciona el sistema desarrollado así como diferentes formas de implementar la solución y los distintos estándares utilizados en el sistema.
- **Propuesta de solución al sistema:** Descripción de los principales componentes, la implementación y diseño del sistema, así como las especificaciones de software y hardware utilizado para cumplir con el objetivo de la tesina.
- **Pruebas y resultados:** Pruebas acerca de viabilidad del proyecto, las configuraciones realizadas y el cumplimiento de los requerimientos planteados al inicio.

- **Conclusiones:** Conclusiones obtenidas del trabajo y los resultados obtenidos, así como observaciones a trabajos futuros que se pueden realizar y omptimización del proyecto.

## Capítulo 2

# Análisis

En esta sección se presentarán los criterios específicos que el sistema de generación musical basado en IA debe cumplir para considerarse exitoso. Estos requerimientos guiarán el diseño y la implementación del modelo y serán la base para evaluar su rendimiento y efectividad.

### 2.1. Requerimientos Funcionales

- **Precisión Melódica:** El modelo deberá ser capaz de producir secuencias melódicas que sean coherentes con el fragmento MIDI inicial, respetando la tonalidad, el contorno melódico y las tendencias estilísticas de la pieza de origen.
- **Armonización Adecuada:** Se espera que las secuencias armónicas generadas complementen las líneas melódicas y mantengan la coherencia con los acordes y progresiones típicas del género musical del fragmento original.
- **Consistencia Rítmica:** El modelo debe mantener la coherencia rítmica, replicando y extendiendo los patrones rítmicos del fragmento inicial de

manera que la pieza generada posea un flujo natural y reconocible

- **Generación Continuada:** Debe ser capaz de generar una pieza musical que fluya lógicamente desde el fragmento MIDI proporcionado, sin interrupciones abruptas o cambios que rompan la continuidad estilística y temática
- **Versatilidad Estilística:** El sistema deberá ser versátil en el manejo de varios estilos musicales, pudiendo adaptar su generación a las características específicas de cada uno, ya sea jazz, clásico, pop, entre otros.

## 2.2. Requerimientos No Funcionales

- **Tiempo de Respuesta:** El modelo deberá generar la música en un tiempo razonable, permitiendo su uso en tiempo real o cerca de este, lo cual es crítico para aplicaciones en vivo o interactivas.
- **Eficiencia Computacional:** La solución propuesta deberá ser eficiente en términos de uso de recursos computacionales, maximizando la calidad de la generación musical sin requerir hardware excesivamente especializado.
- **Capacidad de Entrenamiento:** El sistema debe poder entrenarse efectivamente incluso con conjuntos de datos limitados, aprovechando técnicas como el aprendizaje por transferencia o el aumento de datos para mejorar su rendimiento.

### 2.3. Restricciones Funcionales

Las restricciones del proyecto son limitaciones que podrían afectar el alcance del estudio, la metodología y los resultados finales. Estas restricciones se categorizan en técnicas y de datos.

- Capacidad de Cómputo Disponible: La calidad y la velocidad del modelo de generación musical dependen en gran medida de la capacidad de procesamiento. Los recursos computacionales limitados pueden restringir la complejidad del modelo, la velocidad de entrenamiento y la capacidad de procesamiento en tiempo real.
- Limitaciones del Formato MIDI: Aunque el MIDI es un formato estándar para la música digital, no captura todas las matices de una ejecución musical, como el timbre y la expresión, lo que puede limitar la profundidad y la expresividad de la música generada.
- Duración de Entrenamiento del Modelo: El entrenamiento de modelos de IA, especialmente redes neuronales profundas, puede requerir un tiempo significativo. Restricciones en el tiempo de entrenamiento pueden impactar en la elección de la arquitectura del modelo y en la amplitud de los datos de entrenamiento utilizados.
- Escalabilidad: La capacidad del modelo para escalar y manejar incrementos en el tamaño del conjunto de datos o en la complejidad de las tareas sin una degradación significativa en el rendimiento.

### 2.4. Restricciones de Datos:

- Disponibilidad y Diversidad de los Datos MIDI: La efectividad del modelo está directamente relacionada con la calidad y la diversidad

del conjunto de datos de entrenamiento. Las restricciones en la disponibilidad de datos abiertos y variados pueden afectar la capacidad del modelo para generalizar y producir música en diferentes estilos.

- **Derechos de Autor y Uso de Material Protegido:** La música está sujeta a derechos de autor, y el uso de piezas protegidas para entrenar o evaluar el modelo puede plantear problemas legales. Es necesario asegurarse de que los datos utilizados no infrinjan las leyes de propiedad intelectual.
- **Sesgo de Datos:** Los conjuntos de datos pueden tener sesgos inherentes basados en la selección de música disponible y utilizada para el entrenamiento, lo que podría limitar la capacidad del modelo para generar ciertos tipos de música o reflejar una diversidad de estilos y contextos culturales.

## **2.5. Trabajos relacionados**

En la sección de trabajos relacionados, se efectuará un análisis detallado de investigaciones anteriores y desarrollos tecnológicos en el ámbito de la generación de música con inteligencia artificial. Esta revisión crítica permitirá establecer el estado del arte y destacar las contribuciones del presente estudio. Los elementos a cubrir incluyen:

### **2.5.1. Revisión de la Literatura:**

Se examinarán estudios académicos que abordan la generación de música utilizando IA, incluyendo:

- **Modelos de Redes Neuronales Profundas:** Investigaciones que utilizan RNNs (Redes Neuronales Recurrentes), CNNs (Redes Neuronales

Convolucionales), y GANs (Redes Generativas Adversarias) para la generación de música.

- "DeepBach: a Steerable Model for Bach Chorales Generation" que utiliza modelos de aprendizaje profundo para generar corales en el estilo de Bach.
- Sistemas de Composición Asistida: Herramientas que proporcionan asistencia a compositores y productores musicales, como Google's Magenta o IBM's Watson Beat.
  - Magenta: Music and Art Generation with Machine Intelligence" que presenta una plataforma para la creación de música y arte mediante machine learning.
- Aplicaciones de Modelado de Música: Uso de modelos para tareas específicas como armonización automática, creación de acompañamiento y generación de melodías.
  - Flow Machines: A New Generation of Music Composition Tools, que explora el uso de IA para la composición de música popular.

### **2.5.2. Comparación de Modelos:**

Se evaluarán contra los modelos y sistemas existentes, identificando sus metodologías, capacidades, y limitaciones. Se prestará especial atención a:

- Desempeño en Generación de Música: Cómo los modelos existentes han logrado generar música que es tanto técnicamente sólida como estéticamente agradable.



- Comparación con AIVA, un sistema de inteligencia artificial reconocido por la SACEM (Sociedad de Autores, Compositores y Editores de Música) como un compositor legítimo de música.
- Diversidad Estilística y Adaptabilidad: La habilidad de los modelos para generar música en varios estilos y su adaptabilidad a nuevos géneros.
  - Análisis de "Jukedek", una start-up que desarrolló un sistema capaz de crear música en varios estilos y que fue adquirida por TikTok.
- Evaluaciones de Usuarios externos y Expertos: Las respuestas de músicos, compositores y el público general a las composiciones generadas por IA.
  - Estudios que utilizan encuestas y paneles de evaluación, como Evaluating Music Generation Systems with Steerable Metrics que proponen métricas específicas para la evaluación.