# فصل اول مقدمه

# مقدمه

# ۱.۱. کلیات و پیشینه تحقیق

یادگیری انسان، به عنوان بنیادی‌ترین جنبه‌ی وجودی او، تأثیر عمیقی بر تمامی رفتارهای فعلی و آینده‌ی وی دارد. هر رفتاری که امروز از ما سر می‌زند، نتیجه‌ی مستقیم فرایند یادگیری ما از کودکی تا به امروز است، و در مقابل، مسیرهای یادگیری و رفتارهای آینده‌ی ما نیز ریشه در تجارب و دانش فعلی ما دارند. از این رو، اگر بتوانیم با بهره‌گیری از آموزه‌های گذشته، مسیری بهینه‌تر برای رشد و پرورش خود بیابیم، می‌توانیم به دستاوردهای سریع‌تر و باکیفیت‌تری دست پیدا کنیم [[1]](#OLE_LINK7). این قاعده در تمامی ابعاد زندگی انسان، از جمله یادگیری مهارت‌های تخصصی، صدق می‌کند. یکی از این حوزه‌ها، یادگیری موسیقی و به طور مشخص‌تر، نواختن یک ساز است. شاید در نگاه اول، مدل‌سازی فرایند یادگیری یک ساز با استفاده از الگوریتم‌های برنامه‌نویسی دشوار به نظر برسد، اما با الهام از مفاهیم یادگیری در انسان، این مهم امکان‌پذیر شده است []۲[](#OLE_LINK8).

بدن انسان، به ویژه در حوزه‌ی یادگیری حرکتی (Motor Learning)، دارای مکانیزمی است که با استفاده از تجارب قبلی و وزن‌دهی به آن‌ها، بازتابی از یادگیری را به ما نشان می‌دهد [[3]](#OLE_LINK9). این مکانیزم، دقیقاً همان مفهومی را تداعی می‌کند که امروزه به آن یادگیری عمیق (Deep Learning) می‌گوییم. در واقع، مغز انسان به عنوان یک ماشین پیش‌بینی‌کننده عمل می‌کند و با استخراج الگوهای آماری از سیگنال‌های دریافتی، انتظاراتی را در مورد آنچه در ادامه رخ خواهد داد، می‌سازد. این مدل‌های پیش‌بینی‌کننده به شدت به تجارب قبلی متکی هستند و به همین دلیل، با یادگیری یک سبک خاص، مغز خود را برای پردازش آن سبک "تنظیم" می‌کند [[۴]](#OLE_LINK10).

# ۱.۲. اثر انگشت آهنگساز و یادگیری بهینه

پروژه‌ی حاضر به طور خاص بر یادگیری افراد مبتدی تمرکز دارد. فرض کنید فردی برای اولین بار با یک ساز مشخص، برای مثال پیانو، مواجه می‌شود و پیش از این تجربه‌ی نوازندگی نداشته است. این فرد با توجه به علایق و مهارت‌های قبلی خود، مثلاً داشتن ظرافت در طراحی یا تجربه‌ی بازی‌های ویدئویی سریع، ممکن است دارای ویژگی‌های منحصر به فردی باشد که در فرایند یادگیری موسیقی تأثیرگذار است. برای مثال، فردی با مهارت طراحی ظریف، ممکن است در تمیز دادن نت‌های به هم پیوسته بهتر عمل کند [[5]](#OLE_LINK11). حال، سوال کلیدی اینجاست که چگونه می‌توانیم این ویژگی‌های فردی را در جهت بهبود فرایند یادگیری به کار بگیریم؟

بداهه‌نوازی، مفهومی است که از ترکیب تجارب پیشین و نقطه‌ی شروع هر فرد در مواجهه با ساز شکل می‌گیرد. تفاوت در نقاط شروع افراد، گاهی اوقات به استعداد ذاتی تعبیر می‌شود، اما آیا این مفهوم دقیق است یا صرفاً فرد قطعاتی را برای شروع انتخاب کرده که با ساختار یادگیری و "اثر انگشت" ذهنی او همخوانی بیشتری داشته است؟ [[6]](#OLE_LINK12). ما فرضیه می‌کنیم که اگر یک قطعه‌ی بداهه یا اولیه‌ی انتخاب شده توسط یک مبتدی، از نظر سبک‌شناختی با «اثر انگشت» یک آهنگساز مشخص (شامل الگوهای ملودیک، هارمونیک و ریتمیک) هم‌خوانی داشته باشد، یادگیری قطعات بعدی آن آهنگساز سریع‌تر و موثرتر خواهد بود [[7]](#OLE_LINK13). برای مثال، یک فرد با شیوه‌ی یادگیری متناسب با سبک باروک، اگر با قطعه‌ای از باخ آغاز کند، به دلیل شباهت‌های ساختاری، نورون‌های مغز او با وزن‌دهی کمتر و پایداری بیشتر، بر بهبود حافظه‌ی عضلانی و تکنیک‌های نوازندگی تمرکز خواهند کرد. این امر نه تنها به یادگیری بهتر آن قطعه می‌انجامد، بلکه آمادگی مغز را برای یادگیری قطعات و تکنیک‌های دشوارتر نیز افزایش می‌دهد [[8]](#OLE_LINK14).

# ۱.۳. تفاوت با کارهای مشابه و نوآوری پروژه

پروژه‌ی حاضر با پروژه‌های مشابهی مانند اپلیکیشن «شازم» (Shazam) که به شناسایی آهنگ‌ها بر اساس اثر انگشت صوتی آن‌ها می‌پردازد، تفاوت ماهوی دارد. در شازم، هدف اصلی تطبیق یک قطعه‌ی صوتی با یک پایگاه داده‌ی بزرگ از آهنگ‌های شناخته‌شده برای شناسایی نام آهنگ و هنرمند است [[9]](#OLE_LINK15). این فرایند مبتنی بر «اثر انگشت صوتی» (Audio Fingerprint) است که برای شناسایی دقیق یک قطعه‌ی خاص طراحی شده است.

در مقابل، هدف پروژه‌ی ما شناسایی یک آهنگ خاص نیست، بلکه طبقه‌بندی سبک بداهه‌نوازی یک مبتدی به «اثر انگشت سبکی» یک آهنگساز است [[10]](#OLE_LINK16). «اثر انگشت سبکی» به مجموعه‌ای از ویژگی‌های موسیقایی (مانند ساختارهای هارمونیک، الگوهای ملودیک، و موتیف‌های ریتمیک) اشاره دارد که در تمامی آثار یک آهنگساز بزرگ، مانند باخ یا بتهوون، به صورت پایدار وجود دارد. این ویژگی‌ها در یک مدل یادگیری عمیق استخراج و شناسایی می‌شوند. بنابراین، اگر بداهه‌نوازی شما ویژگی‌هایی مشابه با اثر انگشت سبکی برامس را داشته باشد، سیستم به شما توصیه می‌کند که برای پیشرفت سریع‌تر، از قطعات برامس شروع کنید [[11]](#OLE_LINK17) . این نوآوری در رویکرد، به ما امکان می‌دهد تا به جای شناسایی صرف، یک مسیر یادگیری شخصی‌سازی شده و بهینه برای هر فرد مبتدی ارائه دهیم.

# ۱.۴. بررسی روش‌شناختی و الگوریتم‌های مورد استفاده

برای رسیدن به این هدف، پروژه ما از رویکردهای پیشرفته در حوزه‌ی پردازش سیگنال (Signal Processing) و یادگیری عمیق بهره می‌برد [[12]](#OLE_LINK18) . ابتدا، سیگنال صوتی بداهه‌نوازی توسط الگوریتم‌های پردازش سیگنال به فرمت‌های قابل تحلیل تبدیل می‌شود. در این پروژه، از روش‌های مبتنی بر سیگنال‌های ویولت (Wavelet) و همچنین تبدیل آن‌ها به تصویر (مانند مل اسپکتروگرام (Mel Spectrogram -استفاده شده است.

در نهایت، برای طبقه‌بندی این سیگنال‌ها، از یک مدل یادگیری عمیق پیش‌آموزش‌دیده به نام «رزنت» (ResNet) استفاده شده است که به طور معمول برای طبقه‌بندی تصاویر به کار می‌رود و روی مجموعه داده‌ی«ایمیج‌نت» (ImageNet) آموزش دیده است [[13]](#OLE_LINK19). با تبدیل سیگنال صوتی به مل اسپکتروگرام، ما این مشکل را به یک مسئله‌ی طبقه‌بندی تصویر تبدیل کرده‌ایم. در نهایت، با استفاده از روش «آنسامبل» (Ensemble) نتایج حاصل از تحلیل ویولت و مدل رزنت با یکدیگر ترکیب شده تا دقت طبقه‌بندی نهایی به بالاترین میزان خود برسد [[14]](#OLE_LINK20).

# ۱.۵. ساختار پایان‌نامه

این پایان‌نامه در پنج فصل تدوین شده است. در فصل اول (مقدمه)، کلیات، فرضیه اصلی و نوآوری‌های پژوهش حاضر به تفصیل شرح داده شد. در فصل دوم، مروری بر مبانی نظری یادگیری حرکتی، عصب‌شناسی موسیقی، و رویکردهای یادگیری عمیق در حوزه‌ی تحلیل سیگنال‌های صوتی ارائه خواهد شد [[15]](#OLE_LINK21). فصل سوم به بررسی روش‌شناختی دقیق پروژه و الگوریتم‌های مورد استفاده در پردازش سیگنال و طبقه‌بندی اختصاص دارد. فصل چهارم شامل پیاده‌سازی عملی، تحلیل نتایج، و مقایسه عملکرد مدل‌های مختلف خواهد بود. در نهایت، در فصل پنجم (نتیجه‌گیری)، یافته‌های اصلی پژوهش جمع‌بندی شده، نقاط قوت و ضعف آن مطرح می‌گردد و پیشنهادهایی برای پژوهش‌های آتی ارائه می‌شود.