

موضوع این پایان نامه مدل سازی دست خط فارسی افراد در چارچوب روش های یادگیری ماشین است. مدل های مولد، از نوع احتمالاتی پارامتری، برای این منظور مورد بررسی قرار گرفته اند. هدف اصلی رسیدن به مدلی بوده است که قابلیت سنتز دست نوشته را داشته باشد، یعنی بعد از آموزش دیدن با نمونه های واقعی دست خط یک فرد، بتواند نمونه هایی مصنوعی شبیه به دست خط او را تولید کند. دامنه این تحقیق به خوشنویسی فارسی محدود شده است، اما ایده ها و روش هایی که مطرح می شوند قابلیت تعمیم به مدل سازی دست خط کلی را دارند. همچنین مدل سازی با فرض داده های آفلاین (تصاویر) و در سطح زیرکلمات انجام شده است. از نظر کاربردی، خروجی اصلی تحقیق رویکردی برای تولید خودکار زیرکلمات خوشنویسی شده است، به طوری که شبیه به دست خط یک هنرمند خاص و، همانند دست نوشته طبیعی، دارای تنوع (غیرقطعی) باشند. این رویکرد می تواند جایگزینی طبیعی تر برای نرم افزارها و فونت های خوشنویسی کنونی باشد. ایده کلیدی رویکرد پیشنهادی این است که نوشتن (خوشنویسی) تابع الگوهایی است و دست خط هر فرد در واقع معادل با یک نگاشت (تبدیل) از الگو به نمونه دست نویس است، که البته نگاشتی احتمالاتی (غیرقطعی) است. الگوی هر زیرکلمه به روشی قاعده مند به دست می آید، و در عمل از یک فونت به عنوان الگو استفاده شده است. پس از تطبیق نمونه های دست نویس واقعی هر زیرکلمه با الگوی آن، یادگیری یک نگاشت کلی الگو به نمونه (مشترک برای همه زیرکلمات) توسط یک مدل احتمالاتی مبتنی بر شبکه های عصبی انجام می شود. اصلی ترین پایه های تئوری این کار روش یافتن نقاط متناظر شکل ها، نگاشت دنباله-به-دنباله توسط شبکه عصبی، تعمیم شبکه عصبی به حالت احتمالاتی و شبکه های بیزی هستند، که قسمتی از پایان نامه به مرور و بحث درباره این مباحث اختصاص دارد. ارزیابی مدل پیشنهادی یکی از دشوارترین مراحل تحقیق بوده است. به دلیل نبود مجموعه داده مناسب مسئله، یک مجموعه داده شامل هزاران نمونه زیرکلمه نستعلیق از ۷ خوشنویس گردآوری و برچسب گذاری شد. برای ارزیابی کمی، از روش های عینی متعارف از جمله محاسبه مقدار درست-نمایی استفاده شده است. همچنین عملکرد مدل در مسئله شناسایی نویسنده مورد آزمایش قرار گرفت. اما از آنجا که خروجی سنتز باید از نظر انسان ها مطلوب باشد، به روش های ارزیابی ذهنی نیز نیاز بود. از این رو، پرسش نامه هایی بر اساس ایده آزمون تورینگ برای افراد عادی و کارشناسان خط طراحی شدند. نتایج در مجموع متوسط، اما امیدوارکننده هستند، و نشان می دهند که مدل توانایی تولید خروجی هایی که انسان را از نظر طبیعی بودن فریب دهد دارد، اما برای حفظ ویژگی های منحصر به فرد دست خط افراد هنوز کار بیشتری لازم است.