**پروژه تحقیقاتی: پیش‌بینی قیمت مسکن با استفاده از الگوریتم‌های پیشرفته یادگیری ماشین**

نویسنده: ابوالفضل دانشجو

تاریخ انتشار: خرداد ۱۴۰۴

**چکیده**

در این پژوهش، عملکرد چندین الگوریتم پیشرفته یادگیری ماشین در پیش‌بینی قیمت مسکن مورد ارزیابی قرار گرفته است. با بهره‌گیری از مجموعه‌داده واقعی Boston Housing و اعمال تکنیک‌های مدرن پیش‌پردازش، مدل‌های رگرسیون خطی، درخت تصمیم، جنگل تصادفی، XGBoost و LightGBM پیاده‌سازی و مقایسه شدند. نتایج تحلیل با استفاده از معیارهای آماری استاندارد نشان داد که مدل LightGBM با دقت ۹۰٪ و کمترین میزان خطا (MSE=0.09) عملکرد بهتری نسبت به سایر مدل‌ها دارد.

**مقدمه**

پیش‌بینی قیمت مسکن یکی از چالش‌های مهم در حوزه اقتصاد داده‌محور است. افزایش تقاضا برای تحلیل دقیق بازار املاک، پژوهشگران را به سمت بهره‌گیری از الگوریتم‌های یادگیری ماشین سوق داده است. هدف اصلی این تحقیق، توسعه و ارزیابی مدل‌هایی هوشمند برای پیش‌بینی دقیق‌تر قیمت خانه‌ها بر اساس ویژگی‌های فیزیکی، موقعیتی و جمعیت‌شناختی آن‌ها است.

**روش‌شناسی**

داده‌ها از مجموعه Boston Housing گرفته شده است. پیش‌پردازش شامل حذف داده‌های پرت با z-score، جایگزینی مقادیر گمشده با KNN Imputer و نرمال‌سازی با RobustScaler بوده است. همچنین داده‌ها به‌صورت ۷۰٪ آموزش و ۳۰٪ آزمون با stratified sampling تقسیم شدند. سپس پنج مدل مختلف پیاده‌سازی شد.

**نتایج تجربی**

در جدول و نمودار زیر عملکرد مدل‌های مختلف ارائه شده است. معیارهای MAE، MSE، RMSE و R² برای ارزیابی دقت مدل‌ها استفاده شدند. همان‌طور که مشخص است، LightGBM بهترین عملکرد را داشته است.

**بحث و تحلیل**

مدل‌های Ensemble مانند LightGBM و XGBoost به‌مراتب دقیق‌تر از مدل‌های کلاسیک عمل کردند. LightGBM به دلیل ساختار سبک‌تر و سرعت آموزش بالا، عملکرد بهتری در زمان و دقت نشان داد.

**جمع‌بندی و پیشنهادات آتی**

مدل LightGBM با دقت بالا و خطای کم برای پیش‌بینی قیمت مسکن پیشنهاد می‌شود. برای تحقیقات آینده استفاده از داده‌های موقعیتی، تصاویر ماهواره‌ای و مدل‌های یادگیری عمیق توصیه می‌شود.

