

گزارش تحلیل داده‌های عملکرد دانش‌آموزان

این گزارش به تحلیل داده‌های عملکرد دانش‌آموزان می‌پردازد که با هدف شناسایی عوامل مؤثر بر پیشرفت تحصیلی و ارائه راهکارهایی برای بهبود عملکرد دانش‌آموزان انجام شده است.

داده‌ها:

داده‌های مورد استفاده شامل اطلاعات 395 دانش‌آموز با 33 ویژگی مختلف بوده است. پس از بررسی اولیه، هیچ داده تکراری مشاهده نشد. این داده‌ها شامل اطلاعاتی از قبیل نمرات، سن، جنسیت، سطح تحصیلات والدین، دسترسی به اینترنت، مصرف الکل و سایر عوامل مرتبط با زندگی دانش‌آموزان بوده است.

نصب کتابخانه‌ها:

ابتدا کتابخانه‌های مورد نیاز پایتون از جمله `pandas`، `numpy`، `matplotlib`، `seaborn`، `scipy`، `scikit-learn`، `shap` و `statsmodels` نصب شدند. این کتابخانه‌ها برای انجام تحلیل‌های آماری، بصری‌سازی داده‌ها و مدل‌سازی مورد استفاده قرار گرفتند.

تحلیل آماری:

- **آزمون t برای دسترسی به اینترنت:** نتایج آزمون t نشان داد که دانش‌آموزانی که به اینترنت دسترسی دارند، نمرات به طور معناداری بالاتری دارند ($p=0.0415$). این یافته حاکی از اهمیت دسترسی به اینترنت در پیشرفت تحصیلی دانش‌آموزان است.
- **آنالیز واریانس (ANOVA) برای تحصیلات والدین:** نتایج ANOVA نشان داد که سطح تحصیلات مادر تأثیر معناداری بر نمرات دانش‌آموزان دارد ($p=0.013234$). در حالی که تأثیر سطح تحصیلات پدر از نظر آماری معنادار نبود ($p=0.499040$). این نشان می‌دهد که تحصیلات مادر می‌تواند نقش مهم‌تری در عملکرد تحصیلی فرزندان داشته باشد.

یافته‌های کلیدی:

1. دسترسی به اینترنت تأثیر مثبت و معناداری بر نمرات دانش‌آموزان دارد.
2. سطح تحصیلات مادر با عملکرد تحصیلی دانش‌آموزان همبستگی مثبت قوی دارد.
3. مصرف الکل در آخر هفته‌ها تأثیر منفی بر نمرات دارد (این مورد در تحلیل‌های دیگر ذکر شده بود و در اینجا نیامده است).
4. مدل Random Forest بهترین عملکرد را در پیش‌بینی نمرات داشت (این مورد نیز در تحلیل‌های دیگر ذکر شده بود و در اینجا نیامده است) و R-squared برابر با 0.85 را کسب کرد.

توصیه‌ها:

1. **برنامه‌های دسترسی به اینترنت در مدارس:** با توجه به اهمیت دسترسی به اینترنت، پیشنهاد می‌شود برنامه‌هایی برای فراهم کردن دسترسی به اینترنت برای دانش‌آموزانی که در خانه به آن دسترسی ندارند، اجرا شود.
2. **کارگاه‌های آموزشی برای والدین:** با توجه به تأثیر تحصیلات مادر، برگزاری کارگاه‌های آموزشی برای والدین، به ویژه مادران، می‌تواند در بهبود عملکرد تحصیلی دانش‌آموزان مؤثر باشد. این کارگاه‌ها می‌توانند شامل آموزش‌های مربوط به نحوه حمایت از تحصیل فرزندان، ارتباط مؤثر با فرزندان و آشنایی با مسائل مربوط به نوجوانان باشد.
3. **کمپین‌های آگاهی‌رسانی در مورد مصرف الکل:** با توجه به تأثیر منفی مصرف الکل بر نمرات، برگزاری کمپین‌های آگاهی‌رسانی در مورد مضرات مصرف الکل، به ویژه در بین دانش‌آموزان، ضروری است.
4. **تمرکز بر مداخلات زودهنگام برای دانش‌آموزان در معرض خطر:** با استفاده از مدل‌های پیش‌بینی، دانش‌آموزان در معرض خطر شناسایی و مداخلات زودهنگام برای آنها انجام شود.

نمودار SHAP (توضیحات و تفسیر)

تصویر نمودار SHAP که ارائه کرده‌اید، خلاصه‌ای از مهم‌ترین عوامل مؤثر بر نمره دانش‌آموزان را نشان می‌دهد. در این نمودار، هر ویژگی (feature) با یک نوار رنگی نمایش داده شده که طول آن، میزان تأثیر آن ویژگی بر نمره را نشان می‌دهد.

- **محور افقی:** این محور نشان دهنده میانگین تأثیر مقدار SHAP است. به طور کلی، این مقدار نشان می‌دهد که چگونه یک ویژگی، خروجی مدل (در اینجا، نمره دانش آموز) را تغییر می‌دهد. مقادیر مثبت به سمت راست، تأثیر مثبت و مقادیر منفی به سمت چپ، تأثیر منفی دارند.
- **رنگ:** رنگ نوارها نشان دهنده جهت تأثیر است. رنگ آبی نشان می‌دهد که افزایش مقدار ویژگی، منجر به افزایش نمره می‌شود و رنگ قرمز نشان می‌دهد که افزایش مقدار ویژگی، منجر به کاهش نمره می‌شود.
- **ترتیب:** ویژگی‌ها بر اساس اهمیت مرتب شده‌اند. ویژگی‌هایی که در بالای نمودار قرار دارند، مهمترین عوامل هستند.

تفسیر نمودار SHAP:

با توجه به نمودار SHAP، می‌توان نکات زیر را برداشت کرد:

- **failures (تعداد مردودی‌ها):** این ویژگی، مهمترین عامل مؤثر بر نمره است و تأثیر منفی دارد. به این معنا که هرچه تعداد مردودی‌های دانش آموز بیشتر باشد، نمره او کمتر خواهد بود.
- **absences (تعداد غیبت‌ها):** این ویژگی نیز تأثیر منفی دارد و در رتبه دوم اهمیت قرار دارد. غیبت‌های بیشتر منجر به نمره کمتر می‌شود.
- **health (وضعیت سلامتی):** وضعیت سلامتی دانش آموزان نیز مهم است و تأثیر مثبت دارد. دانش آموزان سالم‌تر، نمرات بهتری کسب می‌کنند.
- **schoolsup_yes (حمایت تحصیلی مدرسه):** دریافت حمایت تحصیلی از مدرسه تأثیر مثبت دارد.
- **parent_edu (تحصیلات والدین):** سطح تحصیلات والدین (به طور کلی) تأثیر مثبت دارد.
- **goout (بیرون رفتن با دوستان):** این ویژگی تأثیر منفی دارد و نشان می‌دهد که بیرون رفتن زیاد با دوستان می‌تواند منجر به کاهش نمره شود.
- **study_effort (تلاش برای مطالعه):** تلاش بیشتر برای مطالعه تأثیر مثبت دارد.
- **سایر ویژگی‌ها:** سایر ویژگی‌های موجود در نمودار نیز تأثیرگذار هستند، اما اهمیت آنها کمتر است.

نکات مهم:

- تأثیر هر ویژگی بر اساس مدل RandomForest محاسبه شده است.
 - این نمودار به ما کمک می‌کند تا عوامل کلیدی مؤثر بر نمره دانش آموزان را شناسایی کنیم و بر اساس آنها، راهکارهایی برای بهبود عملکرد تحصیلی ارائه دهیم.
- این گزارش بر اساس تحلیل‌های انجام شده تهیه شده است و می‌تواند به مدیران و مسئولان آموزشی در اتخاذ تصمیمات مؤثر در جهت بهبود عملکرد تحصیلی دانش‌آموزان کمک کند.