PMP Analysis

Data Analytics Lab

March 2026

Contents

1 Introduction

Reducing reoffending is an important objective within offender management and understanding which interventions generate the greatest impact is essential for effective policy and practice. This study focuses on the Proactive Offender Management Plan (PMP) for the management of offenders to highlight approaches that drive sustained behaviour change and crime reduction.

The primary objectives of this study are:

- To analyse the range of interventions provided under the PMP.
- To assess individuals' characteristics and their influence on reoffending rates and intervention outcomes.
- To evaluate post-intervention reoffending rates.
- To identify evidence-based strategies that most effectively support sustained behaviour change.

The predictive model developed here is for the purposes of highlighting the differences that various features make to the probability of re-offending and is not intended for application to individuals on an on-going basis.

2 Exploratory Data Analysis

Crime data were extracted from Connect - including all recorded offences occurring from April 2021 onwards. Information on PMPs, including their stated objectives, was also collated from Connect for the period after April 2021. In total, nine distinct PMP objective categories were identified. *Attitudes, Thinking and Behaviour* was the most frequently assigned objective, while *Finance, Benefit and Debt* was the least common (**Figure 1**).

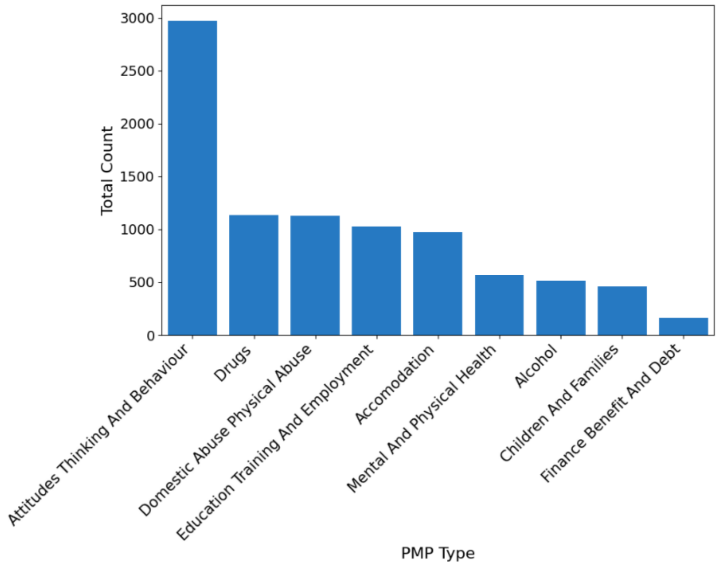

Figure 1. Different types of PMP Intervention Objectives.

Across all recorded offences, fewer than 2% had an associated PMP linked to the incident (**Figure 2**). Possession of weapon offences had the highest proportion of incidents involving individuals who received a PMP, whereas sexual offences had the lowest proportion (0.31%).

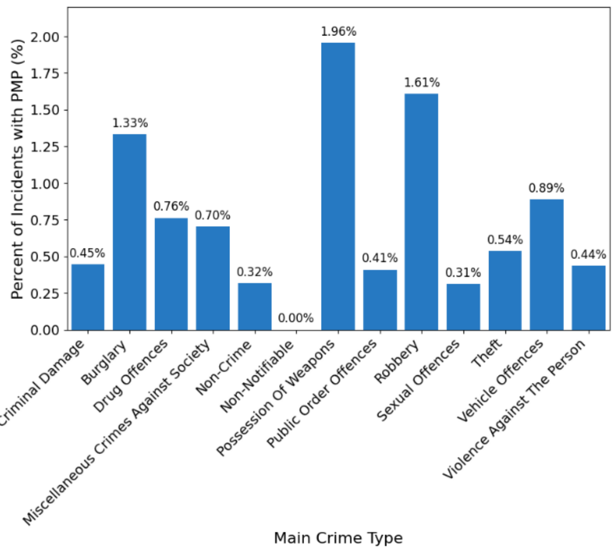

Figure 2. Percent of all crime types that have an associated PMP

Over the last four years, two notable peaks in PMP allocations were observed: the first in early 2022 and the second in 2025 (**Figure 3**). However, a marked decline in the number of PMPs issued has occurred over the most recent year.

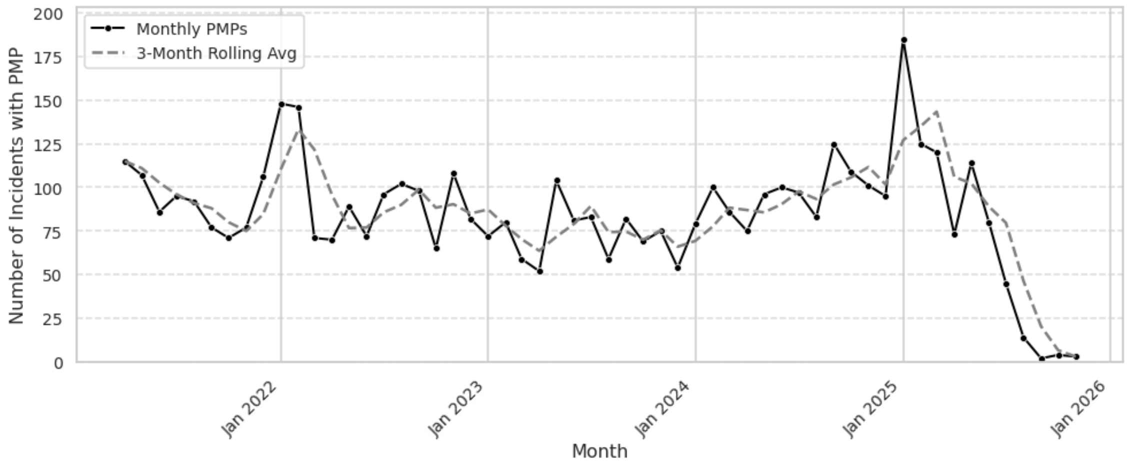

Figure 3: Number of monthly PMPs across all PMP types on connect.

In the majority of cases, PMPs are initiated relatively soon after the initial offence. For the purposes of this study, analysis was restricted to PMPs that commenced after the index offence and within one year of that offence; with each PMP objective being treated as a separate PMP (**Figure 4**). The precise timing of PMP initiation within this one-year window was not explicitly modelled.

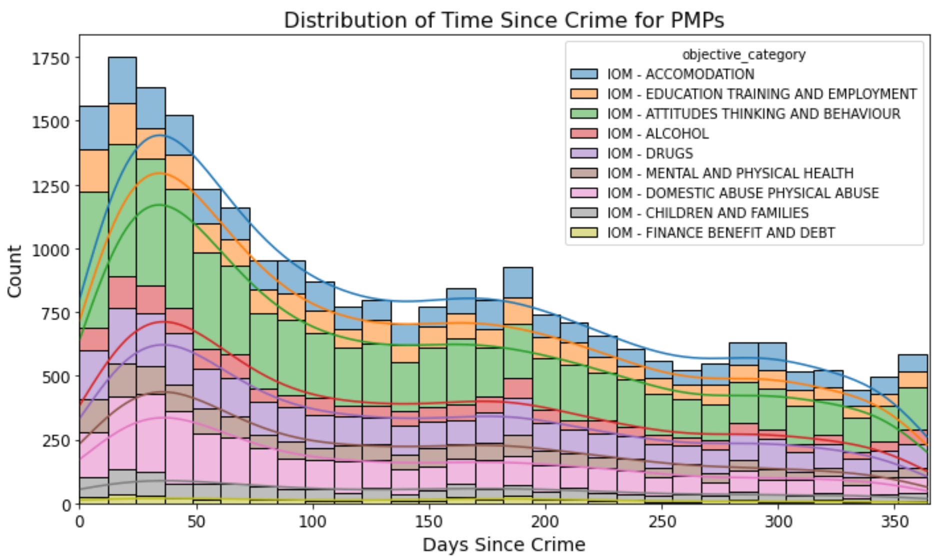

Figure 4: Distribution of PMPs days after initial crime per PMP type.

The study examined the Cambridge Crime Harm Index (CCHI) associated with the index offence, alongside the offence type and subtype, to assess how offence severity relates to reoffending outcomes. To account for prior criminal history, a cumulative CCHI score was calculated for each individual up to (but excluding) the index offence. This measure captures the total harm accrued prior to the offence under study and allows assessment of whether historical offending severity influences reoffending rates.

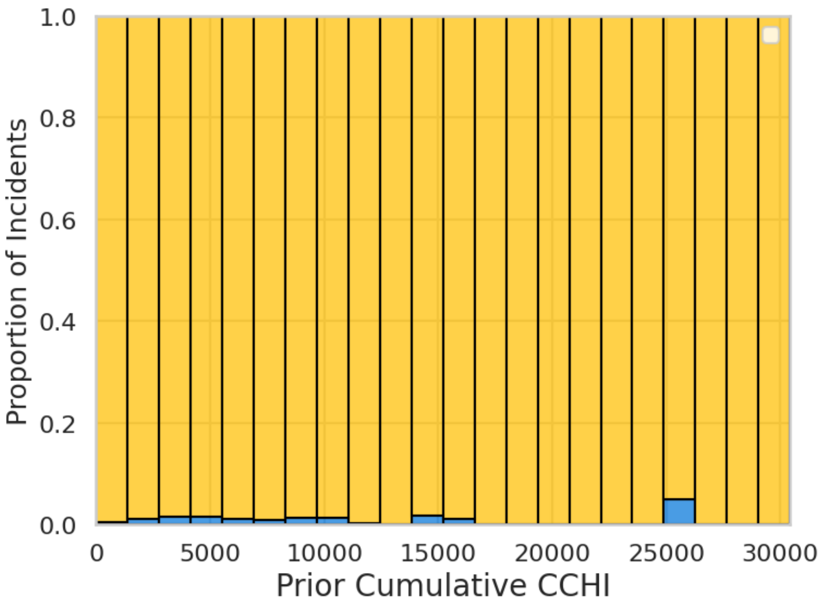

Figure 5: Proportion of cumulative CCHI scores for incidents with PMP (blue) and without PMP (yellow).

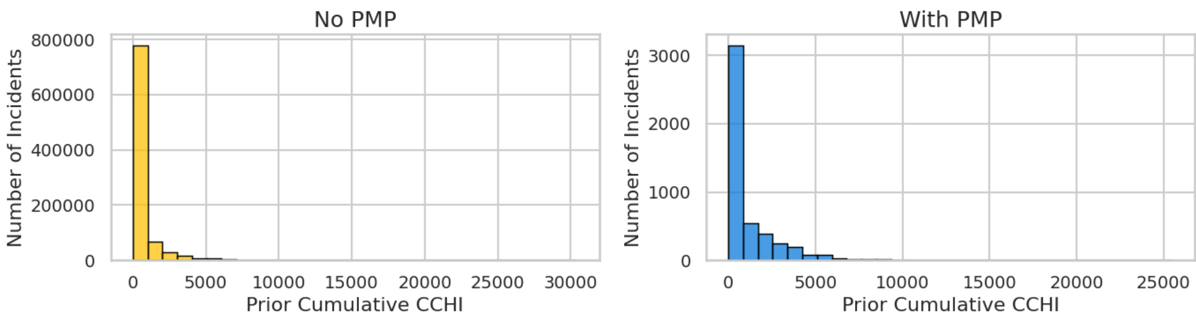

Figure 6: Number of incidents with prior cumulative CCHI scores, with and without a PMP.

The overall distribution of incidents with and without an associated PMP was broadly similar (**Figure 6**). However, incidents involving a PMP showed a higher proportion of individuals with elevated prior cumulative CCHI scores (**Figure 5**). This indicates that PMPs are more commonly allocated to individuals with more extensive or severe offending histories, or to cases assessed as higher risk.

To further characterise prior criminal behaviour, the study incorporated measures of offence escalation. Escalation was defined as an increase in offence severity across consecutive incidents for the same individual, as measured by CCHI. Incidents were ordered chronologically for each individual, and escalation was identified when the CCHI value of a given incident was strictly greater than that of the immediately preceding incident. The first recorded incident for each individual could not be classified as escalatory and was therefore coded as non-escalatory.

When escalation occurred, its magnitude—referred to as the *jump size*—was calculated as the difference between the current incident's CCHI value and that of the previous incident. Where no escalation occurred, either because severity remained constant or decreased, the jump size was set to zero. Decreases in CCHI were not treated as negative jumps, ensuring that the measure captured only upward changes in offence severity.

Escalation frequency was summarised using a cumulative escalation count, which increased by one each time an escalation was observed. This variable represents the total number of escalatory transitions up to and including each incident, rather than a per-incident indicator, and therefore reflects the accumulation of escalatory behaviour over time. In parallel, prior cumulative harm was calculated for each incident as the sum of CCHI values from all preceding incidents, excluding the current offence.

Taken together, these measures distinguish between the occurrence of escalation, the magnitude of severity increases, and the cumulative frequency of escalation across an offending trajectory. This framework enables a nuanced analysis of changes in offence severity over time while maintaining clear temporal ordering.

Reoffending was defined as an individual who was recorded as a suspect or offender of a further offence within one year of the index offence. All PMPs initiated between the index offence and the subsequent offence were included in the analysis.

Socio-economic context was captured using deprivation measures for both the location of the offence and the home location of the individual. Lower Layer Super Output Areas (LSOAs) were derived from the eastings and northings associated with incident locations and individuals' home addresses. These LSOAs were linked to the English Indices of Deprivation (IMD) 2025, which ranks small areas in England across multiple domains including income, employment, health, and education. Lower IMD ranks indicate higher levels of deprivation.

3 Model Selection

Reducing reoffending and assessing the impact of PMP interventions requires modelling of outcomes such as subsequent offending, escalation, and offence severity. The dataset comprises a mixture of numerical and categorical variables, including crime classifications (main group and subgroup), demographic characteristics, geographic identifiers (LPA, beat), prior cumulative Cambridge Crime Harm Index (CCHI), total number of PMPs applied, and escalation measures.

Four separate models were trained against the target variable: logistic regression, XGBoost, Random Forest and CatBoost. CatBoost was selected as the primary modelling framework for this study due to its native support for categorical features. This capability is especially advantageous given the large number of categorical variables present in the data. CatBoost demonstrates strong predictive performance across datasets of varying size, incorporates early stopping to mitigate overfitting, and produces interpretable feature importance measures. It also accommodates mixed data types—numerical, categorical, and Boolean—allowing PMP indicators, cumulative CCHI, and escalation metrics to be incorporated seamlessly. While baseline models such as logistic regression and random forest were built for benchmarking and comparative purposes.

Prior to model training, the dataset underwent a structured preprocessing pipeline to ensure data quality and analytical validity. Irrelevant or redundant variables, including unique identifiers and timestamp fields, were removed to reduce noise and minimise the risk of data leakage. Boolean indicators for individual PMP objectives were retained, alongside derived features capturing intervention exposure. Specifically, an *any PMP* indicator was created to denote whether at least one PMP was applied, and *total PMPs* was calculated as the sum of all PMP indicators per incident, providing a quantitative measure of intervention intensity.

Categorical variables—including crime classifications, demographic attributes (ethnicity, gender, age), and geographic identifiers (LPA, neighbourhood, district, beat)—were preserved in their original form to exploit CatBoost's native handling of categorical data. Numerical features such as prior cumulative CCHI, escalation counts, and total PMPs were retained on their original scale. Missing data were assessed and retained, as CatBoost natively handles missing values in both categorical and numerical features, further simplifying preprocessing.

Feature selection was guided by theoretical relevance and empirical evidence related to reoffending and escalation. Core predictors included prior risk measures (prior cumulative CCHI), incident characteristics (crime type and severity indicators), demographic factors, geographic context, and PMP-related variables (any PMP, total PMPs, and individual PMP categories). This comprehensive feature set captures both offender history and intervention context, providing a strong foundation for modelling.

3.1 CatBoost Implementation

The CatBoost framework was applied to both classification and regression tasks, depending on the defined outcome variable. Binary outcomes, such as escalation occurrence, were modelled using *CatBoostClassifier*, while continuous outcomes, including offence severity and cumulative CCHI, were modelled using *CatBoostRegressor*. Key hyperparameters were selected to balance predictive performance and generalisability, with early stopping employed to prevent overfitting and optimise training efficiency.

All categorical features were explicitly specified to ensure full utilisation of CatBoost's categorical encoding mechanisms. Models were trained and evaluated using a stratified train–test split to preserve outcome distributions, particularly for imbalanced binary targets. Model performance was assessed using appropriate metrics for each task, including Area Under the ROC Curve (AUC) and accuracy for classification models. Feature importance outputs were examined to identify variables most strongly associated with outcomes, supporting both interpretability and actionable insights.

This modelling approach ensures that the analysis delivers not only as accurate predictions of reoffending as possible and escalation but also clear evidence on how interventions, offender characteristics, and contextual factors contribute to post-intervention outcomes. These insights directly support data-driven decision-making and targeted policy development within offender management.

4 Results

Model performance was evaluated using accuracy, ROC AUC, Precision–Recall AUC (PR AUC), F1-score, precision, and recall (**Table 1**).

Across all evaluation metrics, CatBoost demonstrates the strongest overall performance. It achieves the highest accuracy (**0.7656**), indicating superior overall classification capability, and the highest ROC AUC (**0.8409**), reflecting excellent discrimination between reoffending and non-reoffending outcomes.

The PR AUC—particularly important in imbalanced classification problems such as reoffending—is also highest for CatBoost (**0.8760**). This indicates strong performance in identifying high-risk individuals while limiting false positives. The F1-score further confirms CatBoost's effectiveness (**0.8065)**, reflecting a strong balance between precision and recall. The model's recall of **0.8363** shows that a high proportion of individuals who go on to reoffend are correctly identified, a critical requirement for operational decision-making and targeted intervention. However, when comparing the specificities of the models – Logistic regression scores higher (**0.6891**) compared to CatBoost (**0.6660**).

While XGBoost also performs strongly across metrics, it is consistently outperformed by CatBoost. Random Forest offers moderate improvements over Logistic Regression, but both are less effective than gradient boosting approaches in this context.

Table 1: Model results for the prediction of reoffending rates

Model	Accuracy	ROC AUC	PR AUC	F1	Precision	Recall	Specificity
Logistic Regression	0.7031	0.7734	0.8139	0.7372	0.7629	0.7131	**0.6891**
Random Forest	0.7110	0.7757	0.8186	0.7688	0.7212	0.8231	0.5536
XGBoost	0.7293	0.7961	0.8366	0.7808	0.7406	0.8255	0.5944
CatBoost	**0.7656**	**0.8409**	**0.8760**	**0.8065**	**0.7782**	**0.8363**	0.6660

Appendix 1 shows the change in log odds for each of the individual features using CatBoost. Notably all the PMP interventions show low influence on the reoffending rate, whereas ethnicity, age and prior crime count show high feature importance.

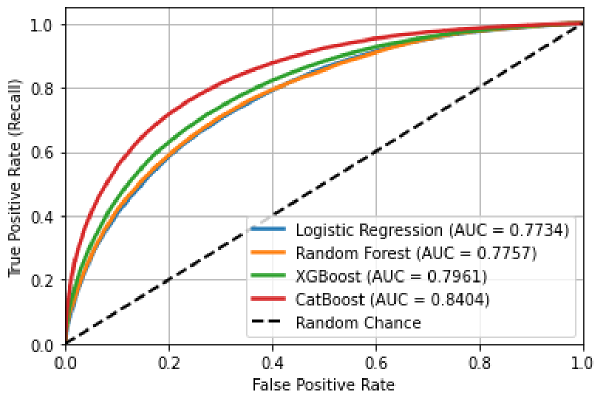

Figure 7: ROC AUC for each of the models

4.1 Without ethnicity

To evaluate the influence of ethnicity on model performance, the models were retrained after excluding ethnicity as a feature.

Table 2: Model results for the prediction of reoffending rates without ethnicity as a feature

Model	Accuracy	ROC AUC	F1	Precision	Recall	Specificity
Logistic Regression	0.6818	0.7527	0.707	0.7648	0.6573	**0.7163**
Random Forest	0.6998	0.7615	0.7601	0.7126	**0.8145**	0.539
XGBoost	0.7123	0.7788	0.7614	0.738	0.7864	0.6083
CatBoost	**0.7444**	**0.8172**	**0.7855**	**0.7701**	0.8016	0.6642

Excluding ethnicity as a feature reduces the overall accuracy and decreases the ROC AUC across all models (**Table 2, Figure 8**). Interestingly, however, omitting ethnicity improves the precision and specificity of logistic regression. In general, most performance metrics decline for all models, including CatBoost. Notably, without ethnicity, Random Forest achieves higher recall than CatBoost—reversing the trend observed when ethnicity was included (albeit at lower levels than the models with ethnicity).

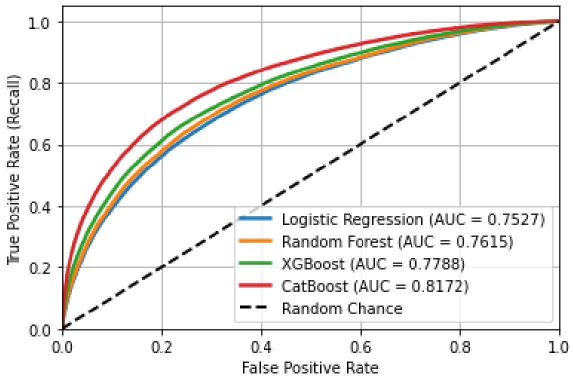

Figure 8: ROC AUC for each of the models without ethnicity as a feature

5 Discussion

This analysis aimed to inform proactive offender management and neighbourhood policing as to the effects of various offender management pathways. The predictive model developed here is for the purposes of highlighting the differences that various features make to the probability of re-offending and is not intended for application to individuals on an on-going basis.

5.1 Precision–Recall Trade-offs and Operational Risk

From an operational perspective, recall is a key metric, as it reflects the model's ability to correctly identify individuals who are likely to reoffend. CatBoost achieves the highest recall among all evaluated models, indicating strong potential for early identification and targeted intervention. Importantly, this performance is achieved without an excessive reduction in precision, ensuring that limited resources are not disproportionately allocated to lower-risk individuals.

5.2 Impact of omitting ethnicity

Omitting ethnicity impacts the performance metrics of the models. When ethnicity is incorporated, most models achieve higher overall accuracy and ROC AUC, suggesting that it provides an informative signal. The specificity of the CatBoost model is very similar with or without ethnicity. However, removing ethnicity can sometimes improve certain metrics, such as precision or specificity in logistic regression. These results highlight that ethnicity can have a nuanced effect on model performance, and its inclusion has been carefully considered in the context of both predictive power and ethical implications (given the needs of the business in relation to this analysis).

5.3 Integration with Neighbourhood Policing

Within the Proactive Offender Management framework, this analysis will support the identification of those PMP pathways that enable the greatest potential for non-recidivism.

5.4 Future Development

Future work could focus on longitudinal validation over time and examine the potential for sensitivity of PMP effects to changing crime patterns.

6 Appendix

6.1 EDA

The full dataset comprised 916,372 samples. Of these, 519,347 were reoffenders, while 397,025 were non-reoffenders. Regarding the PMP interventions, 911,618 samples had no PMPs, and 4,754 samples had at least one PMP. The dataset was divided into training and test sets, with 80% of the data allocated for training and 20% reserved for testing. During the split, the target variable was stratified to maintain the original proportion of positive and negative classes. Additionally, the distribution of the feature was considered to ensure that the samples with and without PMPs were represented proportionally in both the training and test sets.

6.2 Feature importance

Feature	Feature Importance
Ethnicity	20.768
Prior crime count	16.627
Sub sub-group	9.558
Age group	7.266
Home neighbourhood	7.190
Neighbourhood	4.870
Beat	4.692
Prior cumulative CCHI	3.723
Gender	3.638
Main group	3.205
Age	3.183
Home beat	2.759
Sub-group	2.126
Escalated flag	1.012
Escalation count so far	0.932
Prior cat theft	0.916
CCHI of crime	0.853
Prior cat drug offences	0.710
Prior cat non-crime	0.664
Prior cat possession of weapons	0.597
Home IMD rank	0.412
Prior cat arson and criminal damage	0.378
Prior cat public order offences	0.365
PMP attitudes thinking and behaviour	0.345
Home IMD decile	0.337
Prior cat sexual offences	0.325
Prior cat violence against the person	0.318
Home district	0.316
Prior cat robbery	0.254

Home lPA	0.242
District	0.240
Inc IMD rank	0.224
Inc IMD decile	0.187
Prior cat vehicle offences	0.155
Prior cat miscellaneous crimes against society	0.134
Prior cat burglary	0.127
LPA	0.122
Prior cat non-notifiable	0.064
PMP Accommodation	0.052
Escalation jump	0.050
PMP Drugs	0.039
PMP Education training and employment	0.021
PMP Children and families	0.002
PMP Alcohol	0.001
PMP Finance benefit and debt	0.001
PMP Domestic abuse physical abuse	0.000
PMP Mental and physical health	0.000

6.3 Feature importance without ethnicity

Feature	Feature Importance
Prior crime count	21.1795
Sub subgroup	11.0412
Age group	8.7481
Gender	8.4799
Neighbourhood	6.8003
Home beat	6.5257
Prior cumulative CCHI	4.8764
Beat	4.4177
Home neighbourhood	4.2287
Main group	4.0679
Age	3.4447
Subgroup	1.8611
Escalated flag	1.3563
Prior cat theft	1.2988
Escalation counts so far	1.2418
CCHI	1.0510
Prior cat possession of weapons	0.8824
Prior cat non-crime	0.8774
Prior cat drug offences	0.7203
Home IMD rank	0.7137
Prior cat arson and criminal damage	0.6087

Home IMD dec	0.5745
Prior cat public order offences	0.5598
Prior cat violence against the person	0.5461
Prior cat robbery	0.4597
Inc IMD rank	0.4447
Home district	0.4176
Prior cat sexual offences	0.3706
Inc IMD dec	0.3246
Home LPA	0.3164
District	0.2824
PMP Attitudes thinking and behaviour	0.2723
LPA	0.1954
Prior cat burglary	0.1746
Prior cat miscellaneous crimes against society	0.1722
Prior cat vehicle offences	0.1670
Prior cat non-notifiable	0.1125
Escalation jump	0.0839
PMP IOM	0.0571
PMP drugs	0.0208
PMP education training and employment	0.0162
PMP children and families	0.0074
PMP finance benefit and debt	0.0010
PMP alcohol	0.0010
PMP mental and physical health	0.0006
PMP domestic abuse physical abuse	0.0000