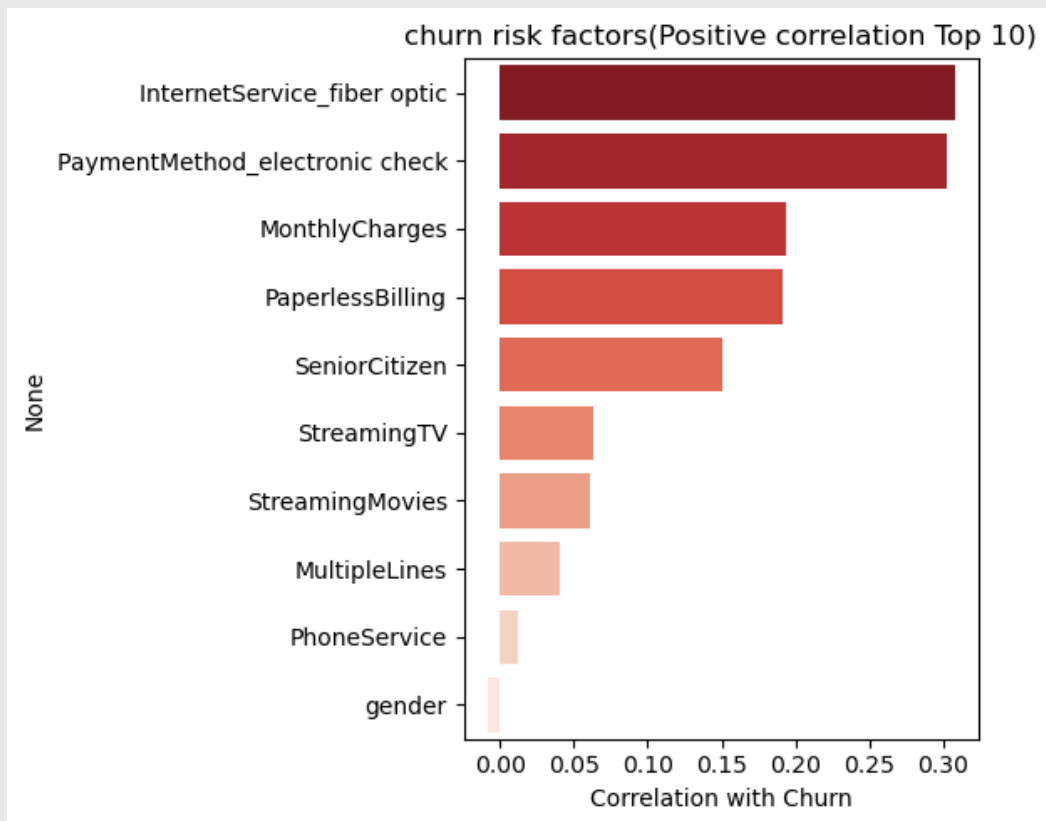


電信客戶流失分析

建模補充

正相關係數圖表

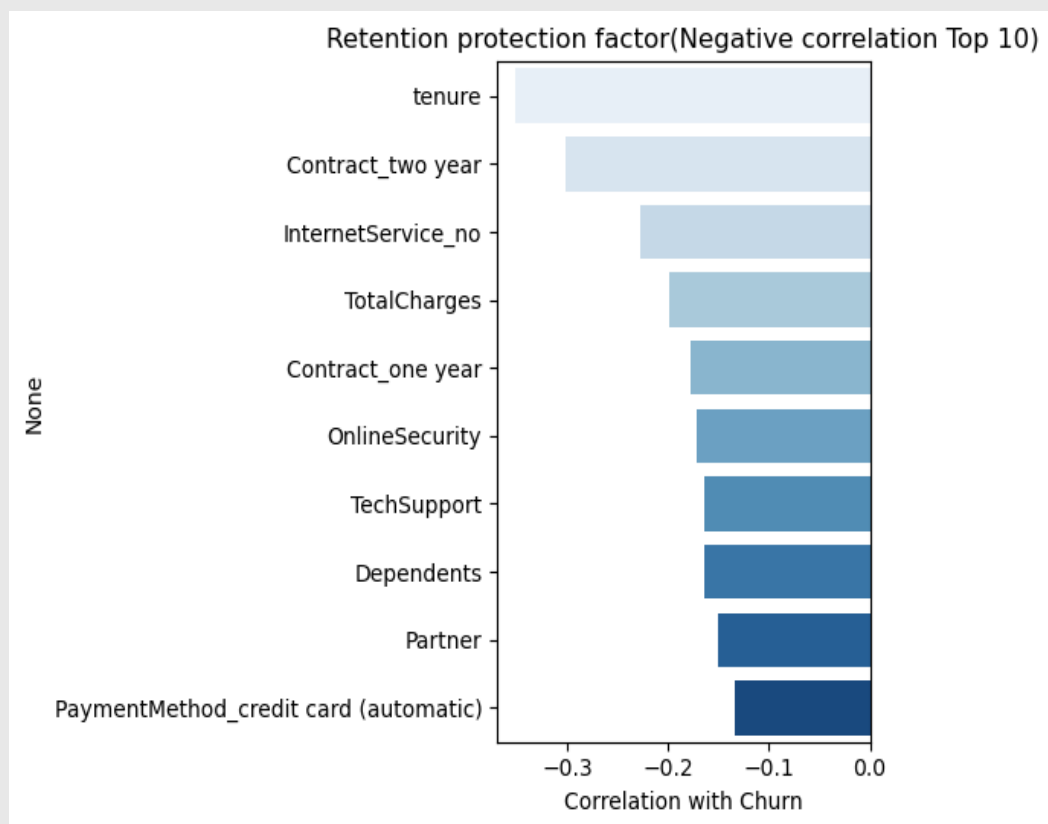


此圖顯示與客戶流失呈正相關的前 10 項因素。其中，以 **光纖網路** 與 **電子支票付款** 的相關性最高 (約 0.3)，代表使用這些服務的客戶更容易流失。此外，**月費** 與 **電子帳單** 也有明顯正相關，表示收費與支付方式影響流失傾向。**高齡族群** 以及 **串流服務 (TV/Movies)**、**多條電話線** 亦顯示出一定程度的正相關。相比之下，**性別** 與 **基本電話服務** 幾乎沒有影響。

◆ 小結論

高費率、光纖服務、電子支付與高齡族群是較明顯的流失風險來源。這與 **EDA 階段的觀察一致**，顯示多種方法下均得到相同結論，提升了分析的可信度。企業應針對這些族群設計差異化方案，以降低流失率。

負相關係數圖表



此圖顯示與客戶流失呈 **負相關** 的前 10 項因素。其中，**使用年資 (tenure)** 的負相關性最強 (約 -0.35)，代表客戶使用時間越長，流失的可能性越低。**兩年合約 (Contract_two year)** 與 **一年合約 (Contract_one year)** 也顯著降低流失風險，顯示合約長度能有效提升留存。

◆ 小結論

影響流失率下降的主要保護因子包含：長期使用年資、較長合約、增值服務、自動付款。這些因素與 **EDA** 階段所得結果一致，顯示 **穩定合約、提高黏著度與提供附加價值服務** 是降低流失率的核心策略。

建模與選擇

	Accuracy Score	F1-score	Recall
Logistic	0.8084	0.60	0.56
KNN	0.7608	0.53	0.53
DecisionTree	0.7921	0.61	0.63
RandomForest	0.7928	0.56	0.50

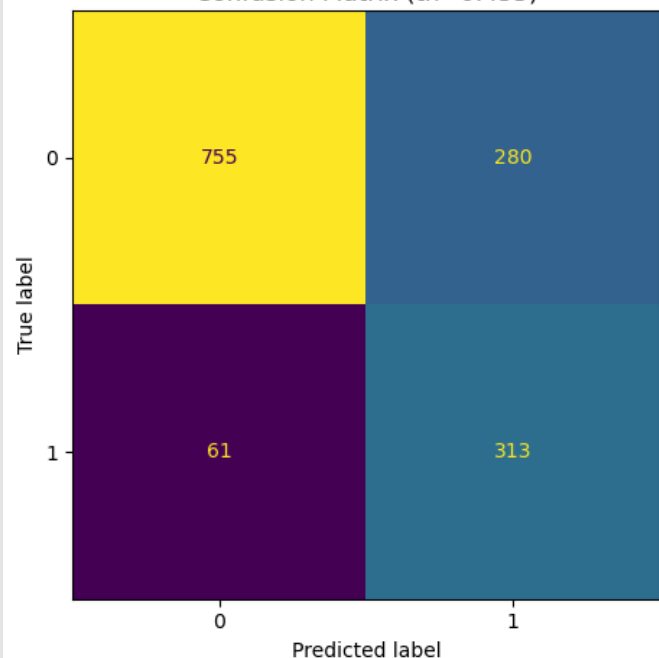
雖然單一決策樹在 Recall 與 F1-score 指標上表現最佳，但考量到其容易過擬合、對新數據的泛化能力有限，因此本作品最終選擇以 **隨機森林** 作為主要優化對象。隨機森林在各項指標雖非最高，但整體表現穩定，且具備較大的調參與優化空間，更適合實務應用。

調參結果

Accuracy Score:

0.7580

Confusion Matrix (th=0.455)

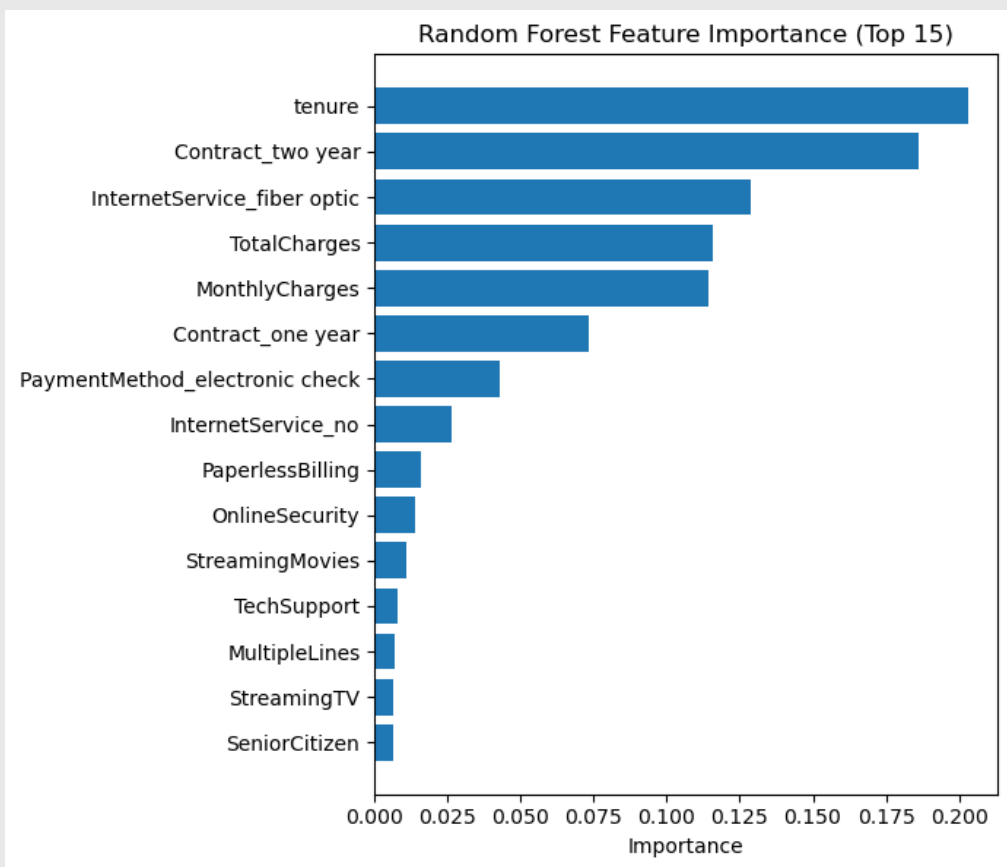


Classification Report

Recall	0.837
F1-score	0.647
precision	0.528

最終以 **隨機森林** 進行調參後，模型在測試集上的 **Accuracy** 為 **0.758**，整體準確度中等，但 **Recall** 提升至 **0.837**，能有效捕捉大部分流失客戶。雖然 **Precision** 僅有 **0.528**，仍存在部分誤判，但高召回率符合「避免錯失潛在流失者」的需求。同時 **F1-score** 為 **0.647**，顯示 **Precision** 與 **Recall** 之間取得一定平衡，整體而言，模型在實務應用中具備相當價值。

特徵值解釋



特徵重要性分析顯示，合約期限（**tenure**、**Contract_two year**、**Contract_one year**）與收費相關變數（**TotalCharges**、**MonthlyCharges**）對流失預測影響最大，凸顯出停留時間與合約型態在判斷是否流失上的關鍵性。其次，網路服務類型（**InternetService_fiber optic**、**InternetService_no**）與付款方式（**PaymentMethod_electronic check**）也具有顯著影響。至於線上安全、技術支援及串流服務等則影響相對較小。整體而言，結果與主作品的**EDA**結論大致相符，進一步強調合約與費用設計在降低流失率上的重要性。

◆ 小結論

1. 合約期限與費用水準是影響流失率的核心因素，顯示延長合約與合理費用設計能有效降低流失風險。
2. 服務型態與付款方式仍具次要影響，與主作品 EDA 結果一致，強化了合約與收費策略的重要性。

結論與展望

- 本研究透過相關係數分析與機器學習建模，驗證了**合約期限、費用水準與付款方式**對於客戶流失的核心影響，並進一步指出光纖網路與電子支付族群為高風險來源。最終模型以**隨機森林**為優化對象，雖在 Precision 上仍有限制，但能**有效提升 Recall**，協助企業在實務上更早鎖定潛在流失客戶。
- 整體結果與主作品的 **EDA 結論大致相同**，進一步強化「**合理費用設計、延長合約與差異化服務方案**」的重要性。未來若能導入更多行為資料並**進行 A/B 測試**，將有助於**提出更具體且可落地的客戶維繫策略**。