Spectrum: Multidisciplinary Journal
Volume 2 Nomor 3, Juli 2025
e-ISSN 3032-5838

PENERAPAN *MACHINE LEARNING* UNTUK MEMPERTAHANKAN LOYALITAS PELANGGAN MENGGUNAKAN *CHURN PREDICTION*

Ardijan Handijono[1*], Zaldy Suhatman[2]
Universitas Pamulang, Tangerang Selatan, Banten, Indonesia
Email: dosen00853@unpam.ac.id, zaldy@unpam.ac.id

ABSTRAK

Penelitian ini membangun model *Churn Prediction* untuk Bank XYZ menggunakan kerangka kerja CRISP-ML dengan data ±700 ribu nasabah dan transaksi tabungan selama enam bulan. Tantangan utama adalah ketidakseimbangan kelas, di mana hanya 1,70% nasabah berstatus *Churn*. Untuk mengatasinya, digunakan teknik SMOTE. Empat algoritma *ensemble learning* diuji, yaitu RandomForest, XGBoost, LightGBM, dan ExtraTrees, dengan *hyperparameter tuning* menggunakan RandomizedSearchCV dan GridSearchCV. Hasil evaluasi menunjukkan LightGBMClassifier memiliki performa terbaik (F1-Score 0,8623; AUC 0,99). Model diimplementasikan berbasis API yang memungkinkan untuk dapat diintegrasikan dengan sistem lain misalnya CRM, dilengkapi pemantauan performa dan *retraining* otomatis. Pendekatan ini terbukti efektif untuk mendukung strategi retensi nasabah melalui prediksi *Churn* yang akurat.

Kata kunci: *Churn Prediction*, CRISP-ML, SMOTE, LightGBM, Bank

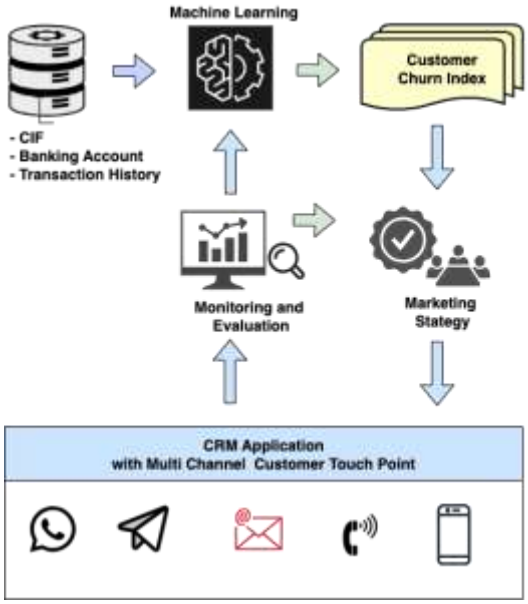
Pendahuluan

Customer Churn terjadi ketika nasabah menghentikan hubungan dengan bank, misalnya melalui penutupan rekening, pencairan deposito, atau pemindahan kredi. Di Indonesia, churn rate nasabah deposito mencapai 19,8% (OJK, 2023), terutama dipicu oleh perbedaan suku bunga yang ditawarkan kompetitor, khususnya bank digital dan fintech (Mayasari & Mahadi, 2023). Nasabah yang sudah *churn* memiliki kemungkinan 50% lebih kecil untuk kembali, sehingga pencegahan lebih efektif dibanding mencari pengganti. Penurunan churn rate sebesar 5% bahkan dapat meningkatkan profitabilitas 25–95% (Otiso, 2024). *Churn Prediction* berbasis AI memungkinkan bank mengidentifikasi nasabah berisiko tinggi berdasarkan pola transaksi, seperti kenaikan signifikan penarikan, penurunan saldo, dan berkurangnya setoran (Verma, 2020). Dengan pendekatan ini, bank dapat mengambil langkah proaktif mempertahankan nasabah dan menjaga profitabilitas. Penelitian ini mengkaji penerapan AI *Churn Prediction* di Bank XYZ serta merancang sistem deteksi dini untuk mendukung strategi pemasaran yang dipersonalisasi Alur proses *Churn Prediction* meliputi beberapa tahap:

1) **Akuisisi dan Integrasi Data** – Tahap awal dimulai dengan pengambilan data nasabah dari *core banking system*. Data yang dikumpulkan mencakup: (a) data demografi (usia, jenis kelamin, domisili), (b) data akun (jenis rekening, saldo rata-rata, status aktif), serta (c) riwayat transaksi (frekuensi, nominal, jenis transaksi). Seluruh data dibersihkan dan diproses sebelum masuk ke tahap berikutnya.
2) **Pemodelan Prediksi Churn** – Data historis digunakan sebagai *input* bagi model *machine learning* yang dilatih untuk mengenali pola perilaku yang

mengindikasikan potensi churn. Model menghasilkan skor churn 0–100, di mana skor lebih tinggi menunjukkan risiko churn yang lebih besar.

3) **Analisis dan Segmentasi Nasabah Berisiko Tinggi** – Hasil prediksi dianalisis untuk menetapkan *threshold* penanganan, misalnya skor ≥ 80. Nasabah dengan skor tersebut masuk ke daftar prioritas retensi.
4) **Penentuan Strategi Retensi** – Tim pemasaran merancang strategi yang relevan, seperti pembebasan biaya administrasi, penghapusan biaya transaksi, penawaran suku bunga kredit lebih rendah, atau peningkatan fasilitas premium. Pemilihan strategi disesuaikan dengan profil dan histori interaksi nasabah.
5) **Monitoring dan Evaluasi** – Efektivitas strategi dicatat dan dievaluasi secara berkala untuk penyempurnaan kebijakan retensi di masa depan.
6) **Personalisasi Saluran Komunikasi** – Pendekatan komunikasi disesuaikan dengan preferensi nasabah, meliputi pesan instan (WhatsApp, Telegram), email, panggilan suara, dan notifikasi *mobile banking*.
7) **Umpan Balik** – Keberhasilan program diukur melalui metrik seperti rasio retensi, peningkatan transaksi, dan respons nasabah. Hasil evaluasi digunakan untuk *retraining* model *machine learning* dan memperbaiki strategi retensi



Gambar- 1 Impementasi *Churn Prediction*

METODE PENELITIAN

Penelitian ini menggunakan kerangka kerja *Cross-Industry Standard Process for Machine Learning* (CRISP-ML) (Handijono, 2022), yang merupakan standar industri dalam pengembangan proyek *machine learning*. Tahapan CRISP-ML meliputi:

1) **Business and Data Understanding** – Menentukan tujuan bisnis, mengumpulkan data yang relevan, serta memahami karakteristik data dan kriteria keberhasilan proyek.
2) **Data Preparation** – Melakukan pembersihan dan transformasi data (normalisasi, *encoding*, pengisian nilai hilang), kemudian membaginya menjadi data latih, validasi, dan uji.
3) **Model Engineering** – Memilih algoritma *machine learning*, melatih model, serta melakukan *hyperparameter tuning* untuk mendapatkan performa terbaik.

4) **Quality Assurance** – Memvalidasi model, mengurangi bias dan *overfitting*, serta mendokumentasikan seluruh proses.
5) **Deployment** – Mengimplementasikan model ke lingkungan produksi, mengintegrasikannya dengan sistem yang ada, dan melakukan pengujian akhir.
6) **Monitoring and Maintenance** – Memantau kinerja model, melakukan *retraining* bila diperlukan, dan memastikan model tetap akurat serta andal.

HASIL DAN PEMBAHASAN

Data mentah diperoleh dari sistem *Core Banking* Bank XYZ, salah satu bank swasta, dengan seluruh informasi rahasia telah disamarkan. Dari ribuan kolom yang tersedia, hanya fitur yang relevan yang diekstraksi, mencakup data sekitar 700 ribu nasabah dan catatan transaksi tabungan selama enam bulan. Variabel yang digunakan meliputi: *customer demographics*, pola transaksi (RFM), skor pinjaman dan koleksi, tabungan dan deposito, penggunaan kanal layanan (cabang, ATM, *mobile banking*, *internet banking*), serta saldo rata-rata. Proses ETL (*Extract–Transform–Load*) sekaligus *data* proses *Machine Lerning*. Tahapan kerangka kerja CRISP-ML meliputi:

1) ***Business and Data Understanding –*** Tujuan model *machine learning* yang dibangun adalah memprediksi kemungkinan nasabah melakukan *churn* atau tidak. Karena menggunakan pendekatan *supervised learning*, diperlukan label pada *training data*. Dalam konteks ini, jumlah nasabah berstatus *churn* biasanya sangat sedikit dibanding total populasi, ini juga membutuhkan penngann khusus.
2) ***Data Preparation*** – Tahap ini mencakup *Exploratory Data Analysis* (EDA) dan *data preprocessing*. EDA dilakukan melalui:
   a. Statistik deskriptif – menghitung rata-rata, median, modus, standar deviasi, dan kuartil tiap variabel.
   b. Visualisasi data – menggunakan histogram, box plot, scatter plot, dan bar chart untuk memahami distribusi, hubungan antar variabel, serta mendeteksi *outlier*.
   c. Identifikasi *missing values* – menemukan data kosong dan menentukan metode penanganannya.
   d. Identifikasi *outlier* – menemukan nilai ekstrem akibat kesalahan input atau kondisi khusus.
   d. Analisis korelasi – mengukur hubungan antar variabel untuk melihat pengaruh satu variabel terhadap yang lain.

   Sedangkan pada proses *data preprocessing* meliputi:
   a. Penanganan *missing values* dan *outlier*
   b. Normalisasi dan standarisasi untuk menyamakan skala variabel numerik
   c. Pengkodean data kategorikal (*one-hot encoding*, *label encoding*) agar dapat diproses oleh algoritma ML
   d. *Feature engineering* untuk menciptakan fitur baru dari fitur yang ada, guna memperkaya informasi bagi model.

Tahap awal terpenting adalah ***labeling***, yaitu menentukan status nasabah (*churn* atau *not churn*) berdasarkan empat indikator (Verma, 2020). Data transaksi enam bulan dipisahkan menjadi dua periode: tiga bulan pertama (*First_3_Months*) dan tiga bulan terakhir (*Last_3_Months*). Perbandingan rasio, kenaikan, atau penurunan nilai pada periode tersebut digunakan untuk menentukan label churn. Hasilnya, seperti terlihat pada Gambar 2, persentase nasabah berstatus churn hanya 1,70%.

Penanganan *Imbalanced Dataset*

Dataset yang digunakan memiliki distribusi kelas yang tidak seimbang (*imbalanced*), di mana nasabah dengan status *churn* (kelas minoritas) hanya berjumlah sekitar 1,70% dari total populasi. Ketidakseimbangan ini berpotensi menimbulkan bias pada model *machine learning*, karena algoritma cenderung memprioritaskan kelas mayoritas sehingga prediksi untuk kelas minoritas menjadi kurang akurat.

Untuk mengatasinya, digunakan teknik SMOTE (*Synthetic Minority Over-sampling Technique*), yang menghasilkan sampel sintetis dari kelas minoritas hingga distribusi kelas menjadi lebih seimbang. Pendekatan ini membantu model mempelajari pola kedua kelas secara optimal, sehingga meningkatkan kemampuan deteksi terhadap nasabah yang berpotensi *churn* (Mqadi et al., 2021).

1. ***Model Engineering* -** Algoritma yang dipilih berasal dari keluarga *ensemble learning* berbasis *Decision Tree*, yang terbukti efektif untuk tugas prediksi *churn*. Model yang digunakan meliputi:

   a. RandomForest – Membangun banyak *decision tree* secara independen dan menggabungkan hasilnya, membuatnya tangguh terhadap *overfitting* serta mampu mengidentifikasi fitur-fitur penting tanpa konfigurasi rumit. **b. XGBoost** – Menghasilkan *tree* secara berurutan, di mana setiap *tree* baru memperbaiki kesalahan *tree* sebelumnya. Algoritma ini sangat kuat dalam menangani data tidak seimbang. **c. LightGBM** – Cocok untuk dataset berukuran besar dengan kebutuhan efisiensi tinggi. Menggunakan teknik khusus untuk membangun *tree* lebih cepat tanpa mengorbankan akurasi secara signifikan. **d. ExtraTrees** – Varian *RandomForest* yang lebih cepat dengan pemilihan *node split* secara acak, mempercepat pelatihan sekaligus membantu mengurangi *overfitting*.

Tabel- 1 Perbandingan Algoritma

Algoritma	Karakteristik Utama	Kelebihan	Kekurangan	Kesesuaian untuk Churn Prediction
Random Forest	Menggabungkan banyak *decision tree* independen	Tahan terhadap *overfitting*, mampu mengidentifikasi fitur penting, relatif mudah digunakan	Waktu pelatihan lebih lama pada dataset besar	Sangat cocok, terutama untuk analisis fitur dan akurasi stabil
XGBoost	*Boosting* berbasis *gradient*; membangun tree berurutan untuk memperbaiki kesalahan	Akurasi tinggi, sangat efektif untuk data tidak seimbang, fleksibel dengan banyak parameter	Membutuhkan *tuning* parameter yang cermat, waktu pelatihan cukup lama	Sangat cocok untuk deteksi *churn* dengan performa optimal
LightGBM	*Boosting* yang dioptimalkan untuk kecepatan dan efisiensi memori	Pelatihan sangat cepat, efisien pada dataset besar, akurasi tinggi	Sensitif terhadap *overfitting* jika parameter tidak diatur tepat	Cocok untuk *churn prediction* pada data besar dan *real-time*
ExtraTrees	Pemilihan *split node* secara acak untuk mempercepat pelatihan	Waktu pelatihan cepat, mengurangi *overfitting*, implementasi sederhana	Akurasi sedikit lebih rendah dibanding *boosting* pada beberapa kasus	Cocok untuk *baseline model* atau saat kecepatan menjadi prioritas

Hyperparameter Tuning

Untuk memperoleh kinerja model yang optimal, diperlukan proses hyperparameter tuning, yaitu pencarian kombinasi nilai hyperparameter terbaik guna mencapai keseimbangan antara bias dan variance, sekaligus menghindari masalah overfitting maupun underfitting. Dalam penelitian ini, hyperparameter tuning untuk RandomForest, XGBoost, dan LightGBM dilakukan menggunakan **RandomizedSearchCV.** Metode ini dipilih karena ketiga algoritma tersebut memiliki jumlah hyperparameter yang cukup banyak, sehingga pencarian secara menyeluruh (*grid search*) akan memerlukan waktu yang sangat lama. RandomizedSearchCV lebih efisien karena hanya menguji kombinasi hyperparameter yang dipilih secara acak dari rentang nilai yang telah ditentukan.

Sementara itu, untuk ExtraTrees digunakan **GridSearchCV**, mengingat algoritma ini memiliki jumlah hyperparameter penting yang lebih sedikit dibandingkan XGBoost atau LightGBM. Dengan GridSearchCV, semua kombinasi yang mungkin akan diuji secara menyeluruh, sehingga dapat diperoleh konfigurasi terbaik secara matematis.

3) Model Validation and Evaluation

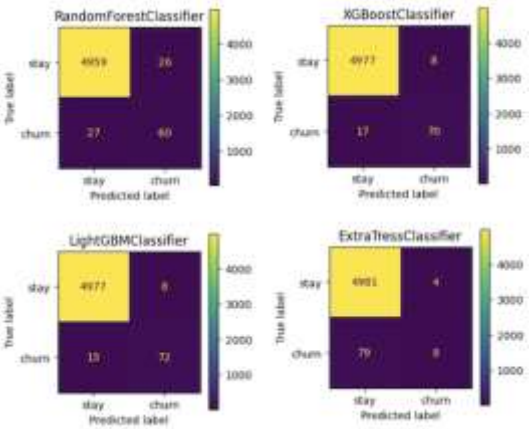
Untuk menilai kemampuan model dalam memprediksi apakah nasabah akan *churn* atau tidak, digunakan beberapa metrik evaluasi berikut:

a) Confusion Matrix – Tabel ini menggambarkan kinerja prediksi model dalam empat kategori:

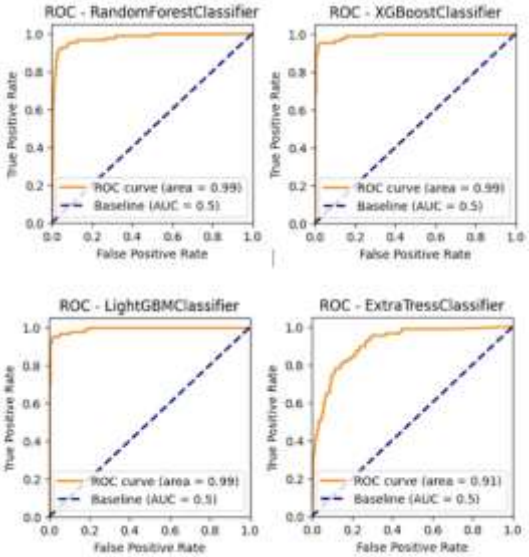
- True Positive (TP): Model memprediksi churn dengan benar.
- True Negative (TN): Model memprediksi tidak churn dengan benar.
- False Positive (FP): Model salah memprediksi churn pada nasabah yang sebenarnya tidak churn.
- False Negative (FN): Model salah memprediksi tidak churn pada nasabah yang seharusnya churn.

b) F1 Score – Merupakan rata-rata harmonis antara *Precision* (tingkat akurasi prediksi positif) dan *Recall* (kemampuan menemukan seluruh kasus positif). Metrik ini sangat penting untuk data yang tidak seimbang seperti kasus churn, di mana jumlah nasabah churn jauh lebih sedikit. Nilainya berada pada rentang 0–1, dan semakin mendekati 1 menunjukkan performa model yang semakin baik.

c) ROC (Receiver Operating Characteristic) & AUC (Area Under Curve) – Kurva ROC menampilkan kinerja model pada berbagai *threshold* klasifikasi, sedangkan AUC menunjukkan luas area di bawah kurva ROC. Nilai AUC juga berada pada skala 0–1, dengan nilai mendekati 1 menandakan kemampuan model yang sangat baik dalam membedakan nasabah churn dan tidak churn.



Gambar- 2 Confution Matrix Chart



Gambar- 3 ROC/AUC Chart

Tabel- 2 Final Score Comparation

Algorithm	F1-Score	ROC Area
RandomForestClassifier	0.693641618	0.99
XGBoostClassifier	0.848484848	0.99
LightGBMClassifier	0.862275449	0.99
ExtratreesClassifier	0.161616162	0.91

Hasil evaluasi menunjukkan bahwa LightGBMClassifier memperoleh nilai F1-Score tertinggi sebesar 0,8623 dengan AUC 0,99. Kombinasi keduanya yang konsisten menjadikan LightGBMClassifier sebagai algoritma terbaik untuk *Churn Prediction* pada studi ini. Berdasarkan hasil evaluasi model, terlihat bahwa **LightGBMClassifier** memperoleh nilai *F1-Score* tertinggi yaitu **0.8623** dengan *ROC Area* sebesar **0.99**. Dengan mempertimbangkan kombinasi *F1-Score* yang tinggi dan *ROC Area* yang konsisten, **LightGBMClassifier** dapat dipilih sebagai algoritma terbaik untuk ***Churn Prediction*** pada studi ini.

5. *Deployment* - Model prediksi churn yang telah divalidasi diimplementasikan ke lingkungan produksi menggunakan arsitektur layanan berbasis API. Integrasi dilakukan dengan sistem CRM atau *data warehouse* perusahaan. Proses deployment

mencakup pembuatan *pipeline* otomatis yang mengekstrak data nasabah terbaru, melakukan *preprocessing* sesuai skema pelatihan, lalu mengirimkan data tersebut ke model untuk menghasilkan skor probabilitas churn. Hasil prediksi disimpan di basis data operasional dan dapat diakses tim retensi melalui *dashboard* analitik.

6. *Monitoring and Maintenance* - Tahap ini memastikan model tetap akurat dan relevan seiring waktu. Pemantauan dilakukan secara rutin pada metrik kinerja utama seperti AUC, F1-Score, dan akurasi pada data terbaru. Selain itu, dilakukan deteksi *data drift* dan *concept drift* untuk mengidentifikasi perubahan distribusi data atau perilaku nasabah yang berpotensi memengaruhi performa model.

Proses pemantauan dapat diotomatisasi menggunakan *model performance dashboard* yang terhubung dengan *pipeline* data produksi. Jika kinerja model turun di bawah ambang batas yang telah ditentukan, prosedur *retraining* akan dijalankan menggunakan kombinasi data historis dan data terkini. Model hasil *retraining* kemudian dievaluasi kembali sebelum digunakan menggantikan model lama di lingkungan produksi.

KESIMPULAN

Penelitian ini berhasil mengimplementasikan *framework* CRISP-ML untuk membangun model prediksi churn nasabah pada Bank XYZ dengan memanfaatkan data transaksi dan profil nasabah. Proses dimulai dari tahap *business and data understanding*, *data preparation*, *model engineering*, hingga *deployment* dan *monitoring*.

Beberapa poin penting yang dapat disimpulkan:

1) Masalah **imbalanced dataset** (hanya 1,70% nasabah churn) dapat diatasi secara efektif menggunakan teknik SMOTE, yang membantu model belajar lebih seimbang terhadap kedua kelas.
2) Dari beberapa algoritma *ensemble learning* yang diuji (RandomForest, XGBoost, LightGBM, dan ExtraTrees), **LightGBMClassifier** menunjukkan performa terbaik dengan nilai F1-Score 0,8623 dan AUC 0,99.
3) Proses *hyperparameter tuning* menggunakan RandomizedSearchCV dan GridSearchCV membantu memperoleh kombinasi parameter optimal yang meningkatkan performa model.
4) Tahap *monitoring and maintenance* memastikan model tetap relevan melalui evaluasi berkala, deteksi *data drift*, serta prosedur *retraining* ketika performa menurun.

Secara keseluruhan, model yang dibangun mampu memberikan prediksi churn dengan tingkat akurasi yang tinggi dan dapat diintegrasikan langsung ke dalam proses bisnis perusahaan.

Saran

1) **Peningkatan Kualitas Data –** Keakuratan model dapat lebih ditingkatkan jika data historis yang digunakan mencakup periode yang lebih panjang serta mencakup variabel perilaku pelanggan yang lebih kaya, seperti interaksi *customer service* atau respon terhadap kampanye pemasaran.
2) **Pengujian Model Lanjutan –** Perlu dilakukan uji coba tambahan dengan algoritma lain seperti CatBoost atau model berbasis *deep learning* untuk melihat potensi peningkatan performa.

3) **Monitoring Otomatis** – Disarankan membangun sistem *automated monitoring* yang tidak hanya melacak metrik kinerja model, tetapi juga mengirimkan notifikasi otomatis jika terdeteksi penurunan performa.
4) **Integrasi dengan Strategi Retensi** – Hasil prediksi *churn* sebaiknya langsung dihubungkan dengan modul *marketing automation* agar perusahaan dapat menjalankan program retensi secara proaktif kepada nasabah yang berisiko tinggi.
5) **Uji A/B di Lingkungan Nyata** – Sebelum model sepenuhnya dioperasikan, sebaiknya dilakukan uji A/B pada kelompok nasabah tertentu untuk mengukur efektivitas model terhadap tingkat retensi aktual.
6) **Kepatuhan terhadap Regulasi Data** – Mengingat data yang digunakan bersifat sensitif, penting untuk memastikan seluruh proses pengolahan dan penyimpanan data mematuhi peraturan perlindungan data yang berlaku.

DAFTAR PUSTAKA

Brito, J. B. G., Bucco, G. B., Heldt, R., Becker, J. L., Silveira, C. S., Luce, F. B., & Anzanello, M. J. (2024). A framework to improve churn prediction performance in retail banking. *Financial Innovation*, *10*(1), 17.

Handijono, A. (2022). Memanfaatkan Data Mining danKnowledge Management System Untuk Mengoptimalkan Strategi Pemasaran pada Bank. *Jurnal Ilmiah Akuntansi Universitas Pamulang*, *5*(1), 268517.

Mayasari, S., & Mahadi, T. (2023, April 5). *Bank digital tawarkan bunga tinggi hingga 9%, OJK angkat bicara.* Https://Keuangan.Kontan.Co.Id/News/Bank-Digital-Tawarkan-Bunga-Tinggi-Hingga-9-Ojk-Angkat-Bicara.

Mqadi, N., Naicker, N., & Adeliyi, T. (2021). A SMOTe based oversampling data-point approach to solving the credit card data imbalance problem in financial fraud detection. *International Journal of Computing and Digital Systems*, *10*(1), 277–286.

Otiso, K. N. (2024). Impact Of Network Quality On Customer Retention; Case Study Of Selected Universities In Kenya, A Customer Relationship Management Approach. *African Journal of Emerging Issues*, *6*(20), 41–50.

Verma, P. (2020). Churn prediction for savings bank customers: A machine learning approach. *Journal of Statistics Applications & Probability*, *9*(3), 535–547.