Prosecution Propensity

Executive Summary

The judicial system in South Africa faces significant challenges due to several factors and not limited to, the high volumes of criminal cases, resource constraints, experience, skills and knowledge, in determining the success of prosecutions, often relying on subjective legal analysis and historical precedent.

In the past five years (2018-2023), SAPS closed over 5.4 million criminal dockets due to insufficient evidence or leads – about 1,051,340 cases in 2022/23 alone. This averages nearly 3,000 cases closed each day without resolution. Meanwhile, the NPA returned about 1.8 million dockets to police for further investigation since 2018, and courts finalise only tens of thousands of cases per year (roughly 74,600 in District Courts in 2022/23), fueling backlogs (over 53,000 cases remain pending).

We propose a prosecution propensity model – an expertise-driven screening model that ingests each case's facts (arrest details, evidence counts, witness info, procedural flags, etc.) and decides whether it is ready for the prosecution procedure. By utilising Merlynn Intelligence Technologies' AI Digital Twin, this model will simulate expert prosecutorial decision-making, assisting prosecutors in making informed decisions regarding case viability, plea bargaining, and trial strategies. Additionally, the model will provide reasons why a prosecution may not be successful, ensuring cases that lack sufficient evidence or other critical factors are not pursued unnecessarily. This will then allow investigating officers to address deficiencies before proceeding further.

We propose that the AI-driven case screening tool (Merlynn's Digital Twin Technology) is integrated at the police-to-prosecution handover (NPA Section 28 referral). This innovation promises to cut wasted effort, speed up prosecutions, and improve fairness by consistently filtering out weak or flawed cases. It aligns with the NPA's push for data-driven innovation and global legal tech trends (AI can automate routine tasks and predict case outcomes).

Early estimates suggest even a 20–30% reduction in futile prosecutions could free tens of thousands of staff-days per year, translating into millions in budget and time savings. By increasing throughput of viable cases and reducing bias from ad hoc decision-making, the model would help bring more offenders to justice efficiently and equitably.

Problem Statement/Business Challenge

Persistent inefficiency and backlog: South Africa's justice pipeline is overwhelmed. Every year, hundreds of thousands of crimes are reported but only a tiny fraction reaches conviction.

Investigations repeatedly stall: In FY2022/23 SAPS closed 1,051,340 cases without result – part of a five-year total of 5.4 million closures due to "lack of evidence or leads". In practice this means each day 3,000 cases end without a single case being enrolled for trial. This represents hundreds of thousands of potentially prosecutable cases that prosecutors must flag as inadequate, wasting detective and prosecutor's time.

Case overload, low throughput: Prosecutors and courts are swamped. The NPA reports extremely high conviction rates (e.g. 95.1% in District Courts, 81.8% in Regional Courts) precisely because only the strongest cases are taken up. Yet even so, courts dispose of only a small portion of the caseload - in 2022/23 roughly 74,600 cases in district courts and 5,500 in regional courts were finalised. Magistrates often sit just 2–3 hours per day on criminal trials, yielding about 0.3 trials finished per day against 2.4 new cases added. This has produced a backlog of over 53,000 outstanding cases. Victims wait years for justice, while many serious offenders roam free.

Resource waste and cost: Each case dismissed or returned after full police work incurs large costs. Detectives may spend weeks or months on a docket that never reaches court. Prosecutors and court officials must review and reject non-viable cases. Overall, the ineffective process diverts billions in public funding where crime already costs about 10% of GDP (roughly R700bn/year). It erodes public trust as the disconnect between crimes reported and offenders punished is a major concern. Bottlenecks like this, thousands of closed dockets per day and overloaded courts, call for a smarter triage solution.

Opportunities and Objectives

Huge caseload: South Africa faces extremely high crime rates where every crime reported enters a labor-intensive pipeline (investigation → prosecution → trial), so inefficiencies have massive aggregate impact.

Lost cases and arrests: In FY2018–2023, SAPS closed 5,446,631 cases due to insufficient evidence. In isolation, 1,051,340 cases were closed in 2022/23. High-volume provinces (Gauteng, WC, KZN) contribute hundreds of thousands of these. Each of these represents wasted investigative effort and unpunished crime. Meanwhile, the

NPA returned roughly 1.8 million dockets to police since 2018 (about 360,000 per year) citing gaps in evidence or procedure. This backlog of referred cases represents rework that drains detective resources.

Selective prosecutions, high conviction rates: The NPA only prosecutes the strongest cases, yielding high convictions (around 95% for District and 82% for Regional cases). But low enrollment numbers indicate many more cases could be heard if proven. The challenge is distinguishing those with lawful arrests, sufficient evidence and sound procedure. Currently, that screening is manual and inconsistent. A data driven model could apply uniform criteria to pre-filter dockets.

Court bottlenecks: Criminal courts handle only a fraction of cases. In 2022/23 District Courts finalised roughly 74 600 matters, Regional Courts roughly 5 500. Submissions outpace decisions by roughly 8 to 1 (e.g. 0.29 trials vs 2.42 enrollments per day). With courts operating just 2–3 hours daily on trials, delays are chronic. Clearing even 10% of current non-viable cases earlier could markedly reduce backlog and remand detention time.

Economic and social cost: Crime has profound costs (roughly 10% GDP), and inefficiency magnifies that. Each delayed or dropped case imposes legal costs (police salaries, admin, detention, etc.). By stopping unlawful or procedurally incorrect cases early, the justice system can reallocate budgets toward effective prosecutions, improve victim confidence, and potentially lessen the overall crime burden.

Proposed Solution/Options

We propose an AI-driven case screening tool (Merlynn's Digital Twin Technology) integrated at the police-to-prosecution handover (NPA Section 28 referral).

The Digital Twin would be trained by experts, using sets of criteria such as the nature of crime, number of complainants/witnesses, evidence quality (e.g. forensic reports), arrest legality flags, suspect criminal history, etc.

Using simulation, the system would output a "prosecution readiness score" for each case. Prosecutors would focus on cases above a threshold score, while lower-scoring cases would be flagged for review or dismissal. Global legal technology has shown that AI can automate routine legal tasks and speed up case assessment.

In practice, our Digital Twin would sit alongside staff (or within) existing case-management systems to help decide whether a case is ready for a prosecutor to pursue. When police submit a docket, the Digital Twin will check data quality and evidence markers with the help of staff to input the data. It then computes the readiness score, and signals whether the docket is ready for prosecution or not. This does not replace legal judgment but augments it: prosecutors can review borderline cases and make final decisions, ensuring the model remains a triage tool.

Explainable AI techniques would be used so that each decision is transparent, helping to guard against bias. Key solution benefits:

- By identifying cases with weak evidence, it prevents time wasted on them downstream, while flagging strong cases (even those that might otherwise slip through) for priority.
- This system aligns with South Africa's innovation initiatives – notably the NPA's call to pilot new approaches in prosecution – and leverages the growing momentum of AI in African legal practice.
- The Digital Twin harnesses decisioning from multiple experts to make the same decision an expert would've made given a certain situation/scenario.

Benefits

Financial (ROI/cost savings)

- **Time and cost savings:** Each avoided prosecution saves resources - fewer detective overtime hours, fewer court scheduling slots, less remand detention time, etc. For example, prosecutors estimate that preparing a case for trial can consume dozens of staff-hours. Cutting caseload by 100 000 cases could save millions of man-hours across SAPS, NPA, and courts. Financially, if one full case costs roughly R10,000 in staff time and overhead, a 10,000-case reduction would save on the order of R100m. Even rough estimates show potential savings in the hundreds of millions of ZAR. Moreover, quicker case screening to determine readiness for prosecution means faster justice for victims and reduced prison/remand costs.

Operational (e.g., increased efficiency, streamlined processes)

- **Improved Case Selection:** Enhances prosecutors' ability to prioritise cases with a higher likelihood of success.
- **Reduced caseload for prosecutors:** If the Digital Twin filters out even 20–30% of cases that would reach the prosecutor, that translates to hundreds of thousands of dockets not requiring full review only for it to be sent back to SAPS/investigators. For context, roughly 360 000 cases per year are currently sent back to SAPS. Saving even 100k docket reviews per year conservatively frees 200 000–300 000 prosecutor-hours (at 2–3 hours of work per case), which can be reallocated to building stronger cases.

Strategic Value (e.g., competitive advantage, improved compliance)

- **Ethical Considerations & Safeguard:** Maintaining transparency in AI-driven decision-making and adhering to legal and ethical guidelines regarding data privacy and security to protect sensitive case information.

- **Increased throughput of sound cases:** With fewer weak cases clogging the system, more court time is available for strong cases. This should boost the absolute number of prosecutions and convictions of serious offenders.

Customer-focused/people-focused (e.g., improved service or satisfaction)

- **Fairness and consistency gains:** The Digital Twin enforces uniform criteria for staff to input the relevant data, reducing arbitrary or biased decisions on which cases proceed. This benefits both the public (through more equal treatment of similar cases) and the accused (by avoiding baseless prosecutions). Victims of well-evidenced crimes can have greater confidence that their cases will be heard if procedures are followed according to the criteria set out. Importantly, the tool would be regularly audited for fairness (e.g. checking that it does not inadvertently disadvantage any group/type of criminal case), in line with best practices in AI fairness.

Investment Rationale

Strategic alignment: This project aligns with government priorities. The NPA's strategic plan explicitly calls for piloting innovative data-driven projects to improve prosecutions. Justice and security policymakers are under pressure to reduce crime costs (10% of GDP currently) and overcrowding, making efficiency upgrades urgent. A funded pilot would demonstrate commitment to the Justice Sector Strategic Plan, which emphasises using technology to raise conviction rates and confidence.

Scale and impact: South Africa's justice system processes very large volumes of cases, meaning a successful implementation of digital twins that automate readiness checks for prosecution will have high leverage. This targets a sizable market- beyond the National Prosecution Authority seeing value in this process, provinces or municipal police could license the model, and the approach could expand to handling all types of cases in all courts.

Cost-effectiveness: Compared to the massive costs of crime (billions of rands per year), court and prosecution delays, the Digital Twin requires relatively modest R&D investment. Over time it yields recurring savings (reduced overtime, less futile case work for prosecutors etc) effectively paying for itself. The long-term ROI is strong, especially as each year's use compounds by accelerating other reforms (like faster police reporting, digital dockets, etc.).

Innovation and tech demand: Legal-technology adoption in Africa is accelerating. This product would be a market-first in South Africa, positioning investors at the forefront of law enforcement AI. It offers both social impact and

financial return as governments may subsidise or co-invest, and private firms (or consultants) can offer deployment and maintenance services. Successful implementation could also generate export potential to other countries facing similar bottlenecks where prosecutors receive cases that are not yet ready to act on.

Public trust and fairness: Demonstrating a fair, data-driven system, it can improve public confidence. With transparent criteria, the justice system can show it is both tough on crime and careful not to over-prosecute. In an era where many citizens feel justice is unequal, a consistent model appeals strongly to both civil society and donors.

Conclusion/Recommendation

By integrating AI technology, specifically Merlynn Intelligence Technologies' AI Digital Twin, into the case evaluation/analyses process, this Prosecution Propensity Model has the potential to revolutionise prosecution decision-making.

The prosecution-propensity digital twin is an innovative, expertise-driven solution to a critical national problem. It leverages existing crime and court data to cut waste and amplify cases that are ready for prosecution, promising measurable time and cost savings. Additionally, by providing reasons for potential prosecution failure, the Digital Twin allows investigators to take corrective action, ultimately contributing to a more effective judicial system.

Backed by official crime facts and proven AI practices, it offers a compelling return - a more efficient, fairer justice system at a fraction of the ongoing cost of current inefficiency. Justice stakeholders gain throughput and credibility, while investors tap into a growing legal-tech market aligned with public-sector needs.