Application of AI Tools in Departure Reason Coding

Concurrence with Virginia’s Guidelines

- ~82% Concurrent

  Sentences fall *within* the recommended ranges of Virginia’s sentencing guidelines.

- ~11% Mitigated

  Sentences are *below* the guidelines — more lenient than recommended.

- < 8% Aggravated

  Sentences are *above* the guidelines — more severe than recommended.

Detailed Departure Reasons are Preserved

- Judges are encouraged to provide **detailed departure reasons** for better insight into sentencing decisions.
- **Original words** from judges are preserved —ensuring authenticity and intent are not lost.
- If judges didn't use the electronic Guidelines system, staff manually **entered departure reasons** into the automated system.

Legacy Process: How Coding Used to Work in Virginia

- **Manual Coding** of Sentencing Departure Reasons
- **Tens of thousands** of sentencing guideline records
- **Up to 3 coded reasons** per departure decision
- **Hundreds of unique codes** accumulated over the years
- **Labor-intensive and repetitive** task
  - Staff involved: 4 full-time & 4 part-time
  - **Accuracy** heavily depended on Individual staff’s knowledge, focus and consistency

Is there Another Way To

- Enhance Operations Without Compromising Privacy
  - ✅ Promotes Automation
  - ✅ Improves Accuracy
  - ✅ Enhances Efficiency
- Keep Human-in-the-loop
  - Staff retains full control of data
- Use In-house resources only
  - No third-party data processing
- To protect data privacy and not use Generative AI

Administrative Challenges

- Restricted AI Use in Virginia’s Judicial Branch
  - Security First: No use of open-source data, Absolutely nothing saved to the cloud
- Human Acceptance Matters:

  “Can a computer truly match the human brain in understanding legal language?”
- Testing the Possibilities:
  - In June 2023, the leadership team at the Virginia Criminal Sentencing Commission decided to begin testing AI solutions —under strict supervision and control
  - Supervised Machine Learning

Re-classify Departure Codes

- **Laying the Foundation for Supervised Machine Learning**
  - **Establish a clear protocol**
    – to guide how departure reasons are processed and categorized.
  - **Re-classify historical departure codes**
    – regroup hundreds of specific codes to more simplified and broader categories.
  - **Reduce dimensionality**
    – streamline data to improve model training and interpretation

Examples

'Aggravating. Offender was the stepparent or foster parent of the victim'
'Aggravating. Offender was a family member or relative of the victim'
'Aggravating. Offender was a family friend of the victim'
'Aggravating. Offender was the supervisor of the victim'
'Aggravating. Offender was the landlord of the victim'
'Aggravating. Offender was a pastor and broke the position of trust'
'Aggravating. Offender broke the position of trust in a work relationship'

Aggravating. Offender broke the position of trust

'Mitigating. Offender has minimal criminal history'
'Mitigating. Offender has only misdemeanor criminal history'
'Mitigating. Offender's first offense'
'Mitigating. First time offender by 18.2-251'
"Mitigating. It is offender's first felony conviction"
'Mitigating. Judge had only juvenile conviction'
'Mitigating. Offender has no felony conviction'
'Mitigating. Judge thought that prior incarcerations are weighted too heavily'
'Mitigating. Judge thought that category I or II and other enhancements to the guidelines are too much'
'Mitigating. Judge thought that convictions occurring after the offense date should not be counted'

Mitigating. Offender has little criminal history or old record

Establish A Database for Supervised Machine Learning

- **Challenge: Non-Standard Language**
  - Accurate training models are **essential** for effective machine learning
  - Judges often use **acronyms and shortened words**
- **Examples**
  - VHSP – Virginia Homeless Solutions Program
  - PRIE – Behavioral Health program
  - Variations of “years old”: YO, Y.O., Y. O., YRS., YR O., Y-OLD
  - Defense abbreviations: D, DEF, DEFS
  - Court/charge terms: CA, CWA, COM. A., CW/A

Collaborating with Staff

- **Shortened list of departure codes developed with staff input**
- AI requires **training to understand human language**
- Humans **review and validate** AI results
- Used **SPSS software** to assign departure codes initially
- Human expertise crucial for:

  Understanding field-specific knowledge

  Decoding acronyms

  Capturing subtle nuances in judges' language
- **Iterative process:**

  Multiple rounds of staff review

  Numerous rounds of program refinement

Python and Text Vectorization

- **Natural Language Processing with Python**
- **Approach:** Vectorizing human text to numerical values
- **Key Concept:** NOT the same as Adobe's summary function
- **Variables:**
  - **Dependent Variable:** 1/0 for departure codes
  - **Independent Variable:** 1/0 for words or phrases in departure reasons
- **Vector Count:** 5000 vectors
  - Higher vector count = More complex calculations

Model Training and Evaluation

- **Data Split:**
  - 80% Training Set and 20% Test Set
- **Techniques Tested:**
  - Logistic Regression
  - Negative Binomial Regression
  - Bayesian Regression
- **Evaluation Metrics:**
  - Accuracy Score, Confusion Matrix (on test data)
- **Outcome:**
  - **Logistic Regression** showed the **highest accuracy (>95%)**
  - Adopted for all dependent variables

Top Ten Departure Codes for Aggravation - Fiscal Year2023

Departure Code for Aggravation	Percentage
Guilty Plea, Agreement, or Joint Recommendation	17.6
No Departure Reason Given for Aggravation	16.7*
Charge was Dismissed, Reduced, or Amended	7.1
Offender had Extensive Criminal History or Repeated Criminal Behavior	6.9
Offender Failed Probation, Alternatives, Court Orders, Appear etc.	5.1
Offense is Cruel, Egregious, or Caused Death	4.8
Offender is Violent or Dangerous (No Children Involved)	4.1
Judge Disagrees With Guidelines	3.8
Victim Impact: Bodily or Emotional	3.5
Children were Involved (Violent)	3.5

* The AI project reviewed all cases with sentences outside the recommended Guidelines range, even if the alternative sentence would be considered within the range. These cases were flagged for further review.

Top Ten Departure Codes for Mitigation - Fiscal Year2023

Departure Code for Mitigation	Percentage
No Departure Reason Given for Mitigation	52.1*
Guilty Plea, Agreement, or Joint Recommendation	15.4
Sentenced to Deferred, Alternative, Treatment, or Service	5.8
Guidelines were Misunderstood or Miscalculated	5.6
Offender is Remorseful	3.5
Offender is Making Progress in Pretrial, Probation, or In General	2
Offender Served, is Serving or to Serve Time on Current or Other Charges	1.8
Evidence Witness Issues, or Old Case	1.8
Guidelines are Modified or Adjusted	1.7
Offender is Cooperative with Law Enforcement	1.6

* The AI project reviewed all cases with sentences outside the recommended Guidelines range, even if the alternative sentence would be considered within the range. These cases were flagged for further review.

LEXIS-AI Alternative Uses

- Expanding Use of LEXIS-AI
  - Initial Use:
    - Identifying **substantially similar offenses** to Virginia laws
  - Exploring New Applications:
    - Summarizing **individual sentences**
  - Limitations:
    - Currently **cannot process large files**
    - Even with future advancements, **human review remains essential**

Key Takeaways

- **Strong leadership is critical**
  - Guiding vision and support are essential for successful AI integration
- **Initial model setup:**
  - Tedious and very time-consuming
  - Required significant effort in data preparation and classification
- **Long-term benefits:**
  - Improved efficiency and accuracy
  - Reduces repetitive manual work
- **Ongoing requirements:**
  - Annual model refinement
  - Continued human review and oversight remain vital

FSR FEDERAL SENTENCING REPORTER

ISSUES FOR AU

RESEARCH ARTICLE | NOVEMBER 01 2025

Virginia's Use of Supervised Machine Learning to Code Judicial Departures from Sentencing Guidelines Recommendations

Catherine Chen; Jody Fridley

Federal Sentencing Reporter (2025) 37 (3-4): 246–249.

https://doi.org/10.1215/10539867-11834102

- **Abstract**: https://read.dukeupress.edu/fsr/article-abstract/37/3-4/246/403307/Virginia-s-Use-of-Supervised-Machine-Learning-to
- **Article** (free-access Author Link): https://read.dukeupress.edu/fsr/article/37/3-4/246/403307/Virginia-s-Use-of-Supervised-Machine-Learning-to?guestAccessKey=7de48130-c8b1-446a-9dd3-d815f5967569

NASC Annual Conference

AI for Sentencing Commissions

August 2025

Elie Alhajjar, PhD
RAND

Opening remarks

AI in criminal justice is not about replacing judgment; it is about giving people *better* information at the *right* time. AI can improve accuracy, fairness, and efficiency across the criminal justice lifecycle if we pair it with strong guardrails, transparency, and human oversight.

Where AI helps right now

- **Data hygiene and integration:** Modern tools de-duplicate records, reconcile aliases and misspellings, and run automated data quality checks. The result is cleaner inputs for guideline studies, recidivism analyses, and racial disparity reviews.
- **Document drafting and review:** Natural language processing (NLP) systems draft memos, summarize case law, check citations, and auto-complete guideline worksheets. This reduces boilerplate work and allows analysts and attorneys to focus on substantive review.
- **Evidence and transcript triage:** Speech-to-text services for calls, hearings, and body-worn camera footage, combined with video analytics, surface relevant segments quickly. Discovery moves faster and the extraction of mitigating and aggravating factors becomes more consistent.
- **Analytic decision support:** Operational dashboards report error rates, calibration, and subgroup disparities for tools in use (e.g., risk screens). These shift debates from anecdote to evidence and inform policy updates and training.

Accuracy: doing the technical homework

- **External validation before deployment:** Models should be tested on local, recent cohorts, with performance reported using confidence intervals rather than single point estimates.

- **Error analysis that matters:** Track false positives and negatives by charge category, supervision level, and key demographics. Make tradeoffs explicit, e.g., how threshold changes affect public safety outcomes and jail populations.
- **Robustness and drift monitoring:** Conduct routine checks for data or population drift, and trigger reviews when distributions shift due to law changes or evolving policing patterns.
- **Explainability for practitioners:** Provide plain-language "why this recommendation?" summaries and list influential factors, while cautioning that correlation does not imply causation.

Fairness: guardrails and governance

- **Define fairness upfront:** Select fairness metrics (e.g., calibration parity, error-rate parity, predictive parity) and acknowledge inherent tradeoffs among them.
- **Disparity testing as a routine:** Perform pre-deployment and post-deployment audits. Report performance by subgroup (race, gender, age, geography) and by interactions (e.g., race × charge).
- **Due process and contestability:** Notify parties when AI-informed assessments are used. Ensure defense access to documentation and a clear path to challenge inputs and outputs.
- **Human-in-the-loop by design:** *AI suggests; people decide*. Require documented rationale when overriding high-impact recommendations (release, detention, departures).

Efficiency: staffing and workflow, not headcount cuts

- **Augment, do not automate away:** Use AI to shrink queues (e.g., transcription, de-identification, form scoring), reduce rework, and reinvest the time saved in interviews, investigation, and quality checks.
- **Role evolution:** Establish cross-functional "Model Steward" teams (legal, data, IT, and ethics) to manage procurement, validation, monitoring, and incident response.
- **Privacy-preserving ways to work:** Employ secure compute enclaves, tiered access controls, and, where appropriate, differential privacy or federated learning to enable inter-agency collaboration without compromising protected data.

Questions leaders should ask

- **Validity:** Where has this model been validated, on what cohorts, and with what performance by subgroup?
- **Explainability:** How will line staff and defendants understand a recommendation?
- **Accountability:** Who owns errors, and what is the remediation path?
- **Monitoring:** What are the drift/disparity triggers and who reviews them?
- **Data governance:** What data are used, who can access them, and how are they retained/audited?

Closing remarks

AI will not deliver justice; people do. But it can help us deliver more consistent, timely, and reviewable justice. Start with low-risk pilots (e.g., data quality, drafting, retroactivity screening), publish validation and fairness reports, and commit to continuous monitoring and stakeholder engagement.