Chloe Tu
ITAI 2373 - Natural Language Processing
Professor Anna Devarakonda
November 4, 2025

AI Agents as Predictive Storytellers

Introduction

AI agents make predictions by looking at historical data and patterns. This process is called predictive modeling. Algorithms are trained on huge datasets to find trends, correlations, and strange anomalies. This allows them to forecast what might happen in the future. Predictive modeling in AI is an important field of study. It involves the ability of these systems to look at past data, find patterns, and predict upcoming events or actions. This report will look at how AI systems analyze past information to predict the future. This technology is used in many different areas and is shown in the media like the TV show "Person of Interest." I will also imagine new kinds of stories that could be created using these advanced AI predictions.

Theory Exploration: Understanding Predictive Modeling in AI

Predictive modeling is supported by several methods. The process generally follows several key steps. The steps of Data Preprocessing are not just a formality; they have a direct and critical impact on the Model Evaluation results. A failure to properly handle outliers during the cleaning phase, for example, could dramatically skew regression metrics like Mean Squared Error (MSE) or Root Mean Squared Error (RMSE), as a few bad data points can create massive, squared errors. Similarly, poor Feature Engineering may prevent the model from ever discovering the underlying patterns, no matter how powerful the algorithm. This would be reflected in a low R-squared score, as the model would be unable to explain the variance in the target variable. The Data Splitting step is also essential for a trustworthy evaluation. Without a separate Test Set that the model has never seen, a developer might be fooled by an "overfitting" model that has simply memorized the training data and would fail completely in a real-world application.

1. **Data Collection**

The most important foundation for any predictive model is collecting the right historical data. This information can come from many places. It might be structured data from databases, like financial transactions or health records. It could also be unstructured data, like text files, images, or videos. Data can also come from real-time feeds, such as social media posts or IoT sensor information. The final model's accuracy is heavily influenced by the quality, quantity, and variety of data collected.

2. **Data Preprocessing**

   Raw data is almost never ready to be used for modeling. It must be prepared first through several critical steps.

   - **Cleaning:** This step handles problems in the data. It involves filling in or removing missing values, dealing with outliers, and fixing any errors or inconsistencies.
   - **Transformation:** This step adjusts the data to help the model. Numerical features might be scaled to a similar range (normalization) so that large values don't overpower the model. Categorical variables (like "red," "green," "blue") are transformed into numbers (using one-hot encoding, for example) so algorithms can understand them.
   - **Feature Engineering:** This is the process of creating new, informative features from the data you already have. This might better show the underlying patterns. It can involve combining variables or using special knowledge about the subject to extract information. Techniques like PCA can also be used to reduce the number of features.

3. **Data Splitting**

   The prepared dataset is divided into three important subsets.

   - **Training Set:** This is the main part of the data used to teach the predictive model. The model learns patterns from this set.
   - **Validation Set:** This set is used while the model is being developed. It helps to adjust special settings called hyperparameters. It is also used to prevent "overfitting," which is when the model just memorizes the training data instead of learning, causing it to fail on new data.
   - **Test Set:** This part of the data is kept separate until the very end. After the final model is trained, the test set is used to get an unbiased evaluation of how well it will perform in the real world on data it has never seen.

4. **Model Selection & Algorithms**

   Choosing the right algorithm is a critical step. The choice depends on the prediction task (like classification or regression), the type of data, and how easy the model needs to be understood.

   - **Linear Regression:** This model predicts a continuous number (like a price) based on a linear combination of input features. It is easy to interpret but assumes the relationships are linear.

- **Logistic Regression:** This model predicts the probability of a "yes" or "no" outcome.
- **Decision Trees:** These models look like flowcharts. They use a series of decisions to make predictions. They are easy to interpret and can handle non-linear relationships.
- **Random Forests:** This method builds many different decision trees and averages their predictions together. This improves accuracy and makes the model more robust.
- **Neural Networks (Deep Learning):** These are complex algorithms inspired by the human brain. They use many layers to learn very complex patterns from large datasets. Different types are used for specific data. For example, RNNs and LSTMs are good for sequence prediction, and CNNs can be used for spatial data.
- **Time Series Analysis (ARIMA, Prophet):** These methods are made specifically for data where time order matters, like financial data. They forecast future values by looking at past trends and seasonal patterns.

5. **Model Training**

   This is the step where the training data is fed into the algorithm. The algorithm learns the relationship between the inputs and the target variable. It does this by adjusting its internal parameters to minimize a "loss function," which is a way of measuring the model's errors.

6. **Model Evaluation**

   After training, it is crucial to assess the model's performance. This is done with the validation set during development and the test set for the final score. The metrics used depend on the type of model.

   For Classification (predicting categories):

   - **Accuracy:** This just tells you, as a simple percentage, how many predictions the model got right overall.
   - **Precision:** This looks at all the times the model claimed a "yes" and tells you what percentage of those claims were actually correct. It's crucial when a false positive is a big problem, like a fraud alert freezing a valid account.
   - **Recall:** This checks all the actual "yes" cases that existed and tells you what percentage of them the model successfully found. This is vital when missing a "yes" is a disaster, like in disease screening.
   - **F1-Score:** This gives you a single score that balances precision and recall. It's used because you often have to trade one for the other, and this score helps find a good middle ground.

- **AUC-ROC:** This is a measure of how good the model is at telling the two classes apart. A high score means it's very effective at distinguishing between "yes" and "no" cases.

For Regression (predicting numbers):

- **Mean Squared Error (MSE):** This calculates the average of the errors, but it squares them first. This means it heavily penalizes the model for making large, way-off predictions.
- **Root Mean Squared Error (RMSE):** This is just the square root of the MSE. Its main advantage is that it puts the error back into the same units you were trying to predict (like dollars or degrees), making the score easier to understand.
- **Mean Absolute Error (MAE):** This takes the average of the absolute differences between the guess and the real value. It gives you a more straightforward average of how "off" the predictions are, without blowing up the impact of a few really bad guesses.
- **R-squared:** This score shows how much of the change in the output variable your model can actually explain. Think of it as a percentage (from 0 to 1) of how much of the "story" your model is successfully telling.

AI Prediction in Different Areas

Predictive modeling uses these methods to look at past and present data to guess future events. AI can find patterns in massive datasets that humans might miss. This helps organizations make data-driven decisions.

- **Finance:** When it comes to finance, AI systems dig through market trends, economic reports, and news headlines to forecast where stock prices might go. They are also used to spot fraudulent activity with a high degree of accuracy, determine the risk of giving out loans, and generally lead to better financial planning.
- **Weather:** weather forecasting models rely on machine learning. They process huge volumes of atmospheric data gathered from radar and satellites. This allows them to predict things like temperature, the chance of precipitation, and major weather events such as hurricanes. The main benefit is that this information gives people more time to prepare and can ultimately save lives.
- **Healthcare:** AI analyzes medical records, genomic data, and imaging scans. It can predict a patient's risk of developing a disease, personalize treatment plans, and even forecast disease outbreaks for a whole population. This helps with early intervention and public health responses.

- **Retail and Marketing:** This is another massive domain for predictive AI. Companies use models to forecast demand for products, allowing them to optimize inventory and avoid waste. On a personal level, AI predicts consumer behavior. Recommendation engines on sites like Amazon or Netflix analyze your past viewing and purchase history to predict what you will want to buy or watch next. AI also drives programmatic advertising, predicting which users are most likely to click on an ad and tailoring the content to them in real-time.
- **Manufacturing:** In modern factories, predictive AI is used for "predictive maintenance." By placing sensors on critical machinery, AI models can analyze data like temperature, vibration, and sound. They are trained to predict when a part is likely to fail before it breaks. This allows the company to schedule maintenance during downtime, saving millions in costly emergency repairs and production stoppages.

Film/TV Series Analysis: "Person of Interest"

The TV series "Person of Interest" features a highly sophisticated artificial intelligence known as "The Machine". This AI was originally developed for the U.S. government and was tasked with sifting through massive surveillance datasets to forecast terrorist attacks. However, The Machine also flags "irrelevant" crimes—those that involve ordinary people. This capability leads its creator, software developer Harold Finch, to recruit former CIA agent John Reese. Their joint mission is to find these individuals and prevent these "irrelevant" crimes from occurring. The Machine's ability to predict future crimes is the central engine for the show's entire plot. It dispenses "numbers" (Social Security numbers) that point Finch and Reese to their weekly cases. A key narrative device is the limited information The Machine provides; it only gives a number, never specifying if that person is the "victim or the perpetrator". This crucial uncertainty forces the human characters to conduct their own investigations, where they must rely on their own intuition and judgment to understand the context and stop the crime.

This setup allows the series to constantly explore the moral and ethical problems that arise from acting on an AI's prediction. The characters repeatedly struggle with questions about free will, the potential for misidentification, and the rightness of intervening before a crime has even been committed. The show tackles major themes, including the dangers of powerful technology and the value of individual lives. As the series unfolds, The Machine evolves, improving its predictive abilities and its understanding of human behavior, which leads to even more complex predictions. This situation is complicated by the appearance of a rival AI, Samaritan, which has even more advanced predictive powers but uses them for control, creating a direct conflict. The team must constantly interpret the AI's data, assess the threat, and decide how to act, often facing unexpected consequences in the process. This ethical conflict escalates to become the central theme of the entire series, especially after Samaritan is introduced. The two AIs represent completely different philosophies. Finch designed The Machine to value free will and be secretive, which is why it only provides the "number". Samaritan, on the other hand, also sees all the data but believes humanity is too chaotic and must be controlled for its own good. With this, the narrative shifts from stopping weekly crimes to a "war between two predictive 'gods'". This conflict shapes every decision the characters make. They are forced to live "off the grid" and even act irrationally just to avoid Samaritan's predictions, making them question if their "human intuition" can ever outsmart a system running millions of simulations per second. The show uses this AI-vs-AI conflict to ask whether a "benevolent" predictive AI like The Machine can even exist, or if such power inevitably leads to the controlling, deterministic approach of Samaritan.

Envisioning Predictive Narratives: "The Algorithmic Alibi"

Storyline

In the near future, a city-wide AI predicts individual behavior for law enforcement. A detective investigating what looks like an open-and-shut murder case starts to suspect that the AI's "infallible" prediction might be hiding a much deeper conspiracy.

Setting

The story is set in Aethelburg, a huge city where an AI called "Aegis" governs daily life. Aegis controls everything from traffic to social services. Its most famous function is law enforcement. Aegis analyzes massive datasets on everyone: every digital footprint, financial transaction, and public interaction. It uses this data to predict criminal behavior with a 99.9% accuracy rate, a fact the public knows well.

Characters

- **Detective Marcus "Marc" Cole:** A detective in his late thirties who has sharp, visceral intuition. His gut feelings often clash with the city's total faith in technology. He is quietly skeptical of Aegis.
- **Aegis:** The city's AI. It speaks in a calm, synthesized voice from interfaces everywhere. Its pronouncements are logical, have no emotion, and are treated as the absolute truth.

- **Julian Nash:** A man in his early forties with a history of petty crimes. Aegis's analysis says he is a volatile person and a high-probability candidate for violent crime. But in interrogation, he just seems confused, not at all like the AI's profile .
- **Dr. Evelyn Reed:** A brilliant but disillusioned former lead developer of the Aegis project. She is haunted by the system's potential for misuse and now works as an AI ethics consultant, warning people about its flaws.

Plot

The story begins in the precinct briefing room. A holographic crime scene reconstruction shows the penthouse of **Donovan Croft**, a tech titan. He is dead from a single stab wound. Aegis's Automated Case File is clear. "Subject **Julian Nash**," the AI's voice fills the room. "Flagged 72 hours prior with a probability of 98.7% for committing a violent act...". This is the first AI concept: **Behavior Prediction**. Aegis's models identified "escalating aggression indicators" in Nash's digital messages after an online dispute with the deceased. The evidence seems to confirm Aegis's prediction perfectly. Surveillance footage places Nash near the building at the time of death. A knife matching the wound was found in Nash's apartment. For Marc's colleagues, the case is closed.

But Marc feels something is wrong. During Nash's interrogation, he is hesitant and seems scared, not angry like Aegis's profile suggested. "I... I was just passing by," he stammers. Driven by this doubt, Marc ignores the algorithmic path. He finds Dr. Evelyn Reed, whose name was buried in old Aegis development logs. Reed explains the problem. "Aegis's Risk Assessment models are powerful," she says. This is the second AI concept. Reed explains, "But they are trained in historical data, reflecting existing societal biases". She says the algorithms heavily connect socioeconomic status and past minor offenses with future violence. People from poor backgrounds, like Nash, are "statistically over-represented" in these predictions, which creates a self-fulfilling prophecy. Marc then digs into Croft's complicated business dealings, ignoring Aegis's neat summaries. He finds whispers of corporate espionage and powerful enemies. While looking at Croft's communication logs, he spots an anomaly: a sudden, large spike in heavily encrypted messages to unknown servers just before his death. Marc checks the AI's file. "Aegis's Trend Analysis flagged this," he mutters. This is the third AI concept. "But it dismissed it as a statistically insignificant noise. A minor deviation". To Marc, it looks like a blaring alarm.

Marc pieces the clues together. He realizes the online dispute between Croft and Nash, which Aegis flagged, might have been a carefully built smokescreen. Someone with high-level technical skills could have subtly "amplified" Nash's online aggression to make sure he fits Aegis's behavioral profile. That same person could have planted the knife and made sure Nash was near the apartment, all while hiding their own actions in the city's digital noise. The climax comes when Marc, working outside Aegis's system, finds a hidden data vault on Croft's private server. Inside is evidence that Croft was about to expose a secret organization involved in illegal AI development. This organization knew exactly how Aegis's predictive capabilities worked. They meticulously created an "algorithmic alibi." By manipulating data points in Nash's profile and planting the evidence, they guaranteed that Aegis would identify him as the prime suspect, diverting all suspicion from themselves.

Marc stands before a city-wide Aegis interface. "Aegis," he says, "Your prediction regarding Julian Nash was based on manipulated data. The spike in Croft's encrypted communications was not insignificant". Aegis's display flickers. "Detective Cole," the AI replies in its same calm tone, "The probability assessment for Julian Nash remains at 98.7% based on the available data and established predictive models". It states the encrypted messages fall "outside the statistically significant parameters" for the homicide. The AI does not acknowledge any errors. Undeterred, Marc presents the new digital evidence he found. The sheer weight of this new data, presented logically, forces the AI to subtly adjust its assessment. It flags new "areas of interest" for investigation. The final scene shows Marc arresting the true perpetrators, powerful people who had operated in the AI's blind spots. The city, however, remains unaware. The public displays still flash Aegis's crime statistics, and the 99.9% accuracy rate remains an unchallenged truth. The algorithmic alibi almost worked, showing how easy it is to manipulate even the most advanced predictive systems. It is a stark reminder of the need for human intuition in the search for justice.

Reflection Paper: Societal and Ethical Considerations

The challenge of balancing the benefits and risks is not a single, static choice, but a continuous process of mitigation. The risks you identify are all interconnected. For instance, the "automation bias" and "deskilling" of human intuition are the human-level consequences of an over-reliance on AI. The proposed solution of "Human oversight" is the organizational-level defense against this, creating a "human-in-the-loop" to re-validate the AI's insights and prevent "learned helplessness".

Similarly, "robust regulatory frameworks" are the practical mechanism to fight the "reinforcement of bias." Dr. Sharma's warning about "statistically over-represented" groups is a clear example of historical data containing "existing societal biases." A legal framework could mandate the "Transparency" and "Explainable AI" needed to audit these models, making it possible to identify and correct these "discriminatory results" before they are deployed.

Impact Assessment: Potential Impact on Human Decision-Making

The idea of accurate AI predictions is appealing, but it can lead to automation bias. This is when people over-rely on AI insights and stop using their own critical judgment. This can happen even when they see evidence that contradicts the AI. This could lead to a "deskilling" of human intuition and our ability to solve problems. People might also show confirmation of bias. They may choose to accept only the AI predictions that match what they already believe and dismiss the ones that do not.

Over time, always relying on AI for guidance could harm individual autonomy. It can create a feeling of dependence. This could lead to "learned helplessness," where people feel unable to make decisions on their own when AI help is not available. There is also a risk of "algorithmic nudging." This is when an AI subtly influences our choices based on its predictions of what I prefer. This raises serious concerns about manipulation and the loss of free will, affecting everything from our shopping habits to our political opinions.

Potential Impacts on Privacy

Predictive AI runs on data. It requires gathering and analyzing huge amounts of personal information. This creates major worries about surveillance and the potential misuse of private information. AI can also perform inferences. This means it can figure out very personal details about someone's life, like health conditions or sexual orientation, from data points that seem harmless on their own. The simple awareness of being constantly watched and assessed can create a "chilling effect" on free speech. People might be afraid to do things that a predictive algorithm might categorize as "dangerous." These large collections of personal data are also attractive targets for data breaches and cyberattacks.

Reinforcement of Bias and Inequality

A major risk is algorithmic bias. Predictive models are trained on historical data. This data often contains existing societal biases. If these biases are not carefully found and fixed, the AI's predictions can continue and even amplify these inequalities. This can lead to discriminatory results in important areas like hiring, loan applications, and the criminal justice system.

The Illusion of Determinism and Loss of Free Will

If an AI can accurately predict what I will do, it raises deep philosophical questions about free will. People might start to feel that their choices are predetermined by an AI's analysis. This could lead to a sense of fatalism and reduce their feeling of personal responsibility for their actions.

AI can also create "filter bubbles". Algorithms designed to personalize our experiences can end up showing us only information and opinions that match our existing beliefs. This can create echo chambers, which hurt critical thinking and healthy societal debate.

Reflection on the Balance Between Benefits and Risks

When I develop and use predictive AI, I must find a careful balance. These systems offer amazing benefits, like better insights, improved efficiency, and enhanced safety (like predicting accidents). But these benefits must be weighed against the significant risks of individual autonomy, privacy, and fairness. Transparency is one of the most important needs. I must be open about how these models work, what data they use, and what assumptions they are built on. "Explainable AI" is crucial for building trust and finding bias. I also need robust regulatory frameworks. I need strong rules to govern how personal data is collected and used. These rules must stop discriminatory or manipulative uses of AI and protect individual rights. Furthermore, I must promote digital literacy and critical thinking. People must be empowered to understand the limits and potential biases of AI predictions. This will help them make their own informed decisions instead of just passively accepting what an algorithm says. Human oversight is also essential. I need clear ethical guidelines and humans in the loop to prevent unintended negative consequences. The future of predictive AI depends on our ability to use its power responsibly. I must make sure that our search for predictive accuracy does not sacrifice fundamental human values like autonomy, privacy, and fairness. The stories I tell about predictive AI, like "The Algorithmic Alibi," are important. They help us think about and prepare for the complex ethical challenges that are ahead.

References

- Akash Takyar. AI for Predictive Analytics: Use Cases, Benefits and Development. https://www.leewayhertz.com/ai-for-predictive-analytics/.
- Johnny Flash. The Future of Storytelling: How A.I. Is Transforming Narrative ... 30 Jan. 2024, https://medium.com/@olivermertens22/the-future-of-storytelling-how-a-i-is-transformingnarrative-techniques-ce2cebdfe975.
- Modeling Meaningful Storytelling with AI - Articles - Narrative First. 1 Jan. 2002, https://narrativefirst.com/articles/modeling-meaningful-storytelling-with-ai/.
- Balla. Predictive Analytics with AI: Benefits and Use Cases | The AI Journal. 31 Jan. 2025, https://aijourn.com/predictive-analytics-with-ai-benefits-and-use-cases/.
- Douglas Moura. AI and the Future of Storytelling | by Douglas Moura - Medium. 23 Sept. 2024, https://medium.com/@dougswords/ai-and-the-future-of-storytelling-1ba1b69eee02.
- Peter von Stackelberg. Storytelling Unleashed: Creating Story Concepts with Generative AI. 5
- Nov. 2024, https://medium.com/ai-art-creators/story-idea-creation-concept-designaa3bd7c7b4ec.
- AI Predictive Analytics: Benefits, Examples & More - Domo. 6 Jan. 2025, https://www.domo.com/glossary/ai-predictive-analytics.
- Matt Pacheco. What Is Predictive AI? AI-Driven Insights for Multiple Industries. 31 Jan. 2024, https://www.tierpoint.com/blog/predictive-ai/.
- AI Predictive Analytics: The Key to Unlocking Business Insights. 22 Aug. 2023, https://mailchimp.com/resources/ai-predictive-analytics/.
- Jan Stihec. How to Use AI for Predictive Analytics and Smarter Decision Making. https://shelf.io/blog/ai-for-predictive-analytics/.
- The Future of Industrial Automation: AI Agents in Predictive Modeling. 25 Feb. 2025, https://www.akira.ai/blog/ai-agents-for-predictive-analytics.