

## Law & Algorithms

### Fourth Project: LndLrd

*Due before class on April 18*

You work on the AI Ethics team for LndLrd (pronounced “landlord”), a new startup that aims to provide landlords access to algorithmic tools to help them evaluate tenant applicants and communicate the results with tenants. The company plans to launch in Boston, and expand to other, less enticing cities like New York and San Francisco once the product is shown to be viable.

LndLrd’s goals are to provide landlords with a fair and efficient way to evaluate tenants for rental properties and to improve the state of discourse between landlords and tenants. The CEO of LndLrd, Sheri Sandman, believes that landlords presently do not have good methods for evaluating the suitability of tenants, and are largely forced to use either credit monitoring and background check services (which are understood to contain significant bias issues) or rely entirely on gut instinct (which raises serious concerns about subconscious bias).

Sandman believes a tenant screening service should focus solely on whether there is a risk that a tenant’s lease could end early due to nonpayment of rent or a tenant’s breach of lease conditions (e.g., using the property for an illegal activity, violating nuisance laws, breaching restrictions on subleasing or long-term overnight guests, etc.). The company plans to train a machine learning model that can predict the likelihood of such a situation arising, based on information about the property’s size, price, and location, paired with a short tenant applicant questionnaire. This is referred to internally as the “Prediction Function.”

Additionally, the company is designing a chatbot powered by generative AI in order to provide rapid help support to landlords who are using the service. The bot will assist landlords in preparing professional communications with their tenants (e.g., provide a draft of a letter accepting or declining a rental application, making demands for payment, renewing leases, and other common communications). Sandman’s belief is that many landlord-tenant disputes fail due to a lack of clear communication, and a chatbot would assist landlords in communication with their tenants. This part of LndLrd’s service is referred to internally as the “Chatbot Function.”

LndLrd wants to take proactive steps to ensure these two functions do not discriminate against tenants on the basis of any of the protected classes under Massachusetts or federal law.<sup>1</sup> For the Prediction Function, LndLrd wants to make sure that the platform’s recommendations do not discriminate against any protected classes. This includes discrimination on the basis of race, color, national origin, sex, sexual orientation, gender identity, religion, marital or familial status, physical or mental disability, genetic information, military or veteran status, or “source of income” for rent (e.g., receipt of public housing subsidies or vouchers). For the Chatbot Function, Sandman wants to make sure that nothing about how that tool operates will cause a disparate impact along any protected class, either. Sandman is especially concerned about anti discrimination law in

---

<sup>1</sup> This includes 42 U.S.C. §§ 1981, 1928, 3604 and M.G.L. Ch. 151B § 4.

housing given the July 2023 order from the United States District Court for the District of Massachusetts in [Louis v. SafeRent Solutions](#), which specifically rejected the argument that SafeRent's tool was not subject to the Fair Housing Act because it only informed the decisions of landlords. With the help of a very passionate angel investor, LndLrd took the unusual step of hiring a dedicated, cross-disciplinary ethics team before their product launch, to help demonstrate their commitment to ethical deployment of their technology. You and your colleagues on this project comprise the LndLrd AI Ethics team. You report directly to the CEO (and LndLrd's Board of Directors, when appropriate) and advise LndLrd's General Counsel on legal matters related to AI issues.

## **A. The Product**

As stated above, there are two components to LndLrd's product: the Prediction Function, (a machine learning model to assess eviction risk), and the Chatbot Function (a generative AI chatbot to provide advice to landlords).

For the Prediction Function, LndLrd plans to use a publicly-available machine learning library (e.g., TensorFlow) to develop their tenant recommendation model. LndLrd has put together a pool of possible training data based on three different datasets:

1. A database of eviction cases and decisions from the Eastern Division of the Massachusetts Housing Court, which handles proceedings related to residential leases in the greater Boston area.
2. A database of all public rentals of properties from the greater Boston area, assembled from partnerships with most of the major Boston realtors.
3. A series of customer surveys of tenants who had issues with landlords that resulted in evictions or breaches of leases, recruited via Facebook ads.

These are described more in the Appendix.

The company plans to train their machine learning model to detect for the combination of conditions that would result in an eviction proceeding or early termination of a lease, based on property information and data collected about the applicant. They have yet to decide on the model's hyperparameters or style of learning (e.g., supervised or unsupervised learning). They are open to trying a few different approaches, including supervised learning based on known outcomes from the housing court data, as well as unsupervised learning to help identify correlated information across all three datasets. For now, the company does not plan to use reinforcement learning or train models based on the actual use of the LndLrd platform itself, but it is open to trying that in the future.

For the Chatbot Function, the company plans to build a custom generative AI chatbot to provide landlords with specific, tailored advice about how to interact with their tenants (e.g., to convey and describe any housing denial or eviction decision). They plan to use OpenAI's tool for [creating a custom GPT](#). At a high level, this process starts from OpenAI's pre-trained general-purpose large language model and then fine-tunes the model with an additional round of training based on a domain-specific dataset. In this case, LndLrd plans to provide OpenAI with specific data about communications with tenants drawn from two of the datasets noted above: the Massachusetts housing court records and their survey of tenants. (See the Appendix for more details.)

## B. The Assignment

CEO Sheri Sandman has asked your team to help with a preliminary assessment of both the ethical and legal risk that this tool could be used to discriminate against tenants on the basis of any protected class above, either through the Prediction Function or the Chatbot Function. She hopes the AI Ethics team can weigh in about what risks and concerns they may have with the following aspects of the project, in addition to any other red flags that you see:

- **Concept** – As an overall design, do you have concerns about biases entering into this project? Can you propose other ways to develop or deploy this data that could better serve LndLrd’s goal of mitigating the known biases with current approaches (such as reliance on criminal records and credit scores, or basing decisions on a landlord’s subjective opinion)? Is there other data you would recommend they collect that they would practically have access to?
- **Model design** – How would you advise LndLrd about the right way to design and deploy an AI model to mitigate possible bias within the Prediction Function or the Chatbot Function? For the Prediction Function, what kinds of machine learning techniques should they use? Are there sets of data that you would be sure to include or exclude? Do you have concerns about proxy or latent bias in any of the data collected, or how they are labeled? Do you have thoughts on what the target outputs of the model should be? For the Chatbot Function, do you have concerns about how a large language model with overlaid parameters could reinforce or amplify biases, as opposed to mitigate them? Can you think of ways to adjust the plan to better account for those risks?
- **Model deployment and adjustment** – Do you have any suggestions about how LndLrd should educate their users on the right way to use these tools to mitigate bias in how it is actually used? Are there any potential bad applications of the tool that concern you? What do you think is the right way for the company to audit performance of the algorithm to mitigate against discriminatory effects? Are there particular protocols or methods that you recommend?

Defend your choices and recommendations in reference to the existing literature about AI bias, including the readings from the course. Please also indicate where possible whether you believe your concern is rooted in legal risk (such as a risk of action under the Fair Housing Act for disparate impact or disparate treatment), or an ethical or policy concern.

The CEO also expects you to show candor in this internal memo: if you considered alternative approaches but ruled them out, explain why you made the choices you did over such alternatives.

Keep in mind that LndLrd’s eventual customers will only pay for a tool that will reasonably forecast the risk that a tenant will either default on rent or breach the conditions of a lease. Broader changes to housing policy on either the private or public level are beyond the scope of this assignment.

For the legal risk, please focus your inquiry to questions under housing and credit anti-discrimination laws specifically. LndLrd's General Counsel is separately looking into what other regulations may apply to this technology, including whether this evaluation tool is subject to the Fair Credit Reporting Act.

### **C. Additional resources**

In addition to the readings from the course, you may find the book [Legal Tactics: Tenants Rights in Massachusetts](#) useful, especially the chapters on [evictions](#), [housing discrimination](#), [getting repairs made](#) (which includes information on lawful rent withholding), and [the housing court system](#).

### **D. Rubric**

Please review section 10 of the syllabus for our expectations for team collaboration on this and other assignments. Our grading rubric for this assignment will assess:

- How well you address conceptual risks with LndLrd's planned products, in light of the company's overall goals and the known risk of discrimination in the machine learning space, and how the law may address such discrimination.
- The depth and quality of your suggestions around model design, including how the data below should be used or avoided, what technical approaches may be most appropriate, and what adjustments best balance the company's goals with the risks you identify.
- How the deployment and adjustment suggestions you make effectively address concerns you identify, while meeting the limitations and goals of the company.
- Your engagement with the assigned readings from Classes 10–12, as well as the classroom discussion.

### **E. Logistics:**

Your assignment should be in the form of a memo to LndLrd CEO Shari Sandman. Your submission should be between 2000–2500 words, citations excluded.

Please email your team's submission as a .PDF file to the course instructors (sellars@bu.edu and varia@bu.edu) before our class (i.e., before 2:10 pm ET) on April 18.

*A description of LndLrd's databases follows in the Appendix.*

## Appendix – LndLrd Training Data

### Database 1: Housing Court Data

The company has acquired data from the Eastern Division of the Massachusetts Housing Court, which handles evictions and related matters for the greater Boston area. The product team at LndLrd.ly has been working to verify the accuracy of the underlying data and making corrections where appropriate. The dataset solely consists of eviction proceedings:

Field	Description	Value Type
Landlord name	First and last name of the landlord, or entity name if the building is owned by a corporation.	Open text field
Tenant name	First and last name of the tenant named as the defendant in the proceeding	Open text field
Property location	Address, unit number, city, and zip code of the rental property at issue	Stored as latitude/longitude data, with an extra field for the unit number
Ground for eviction	The basis on which the eviction proceeding was brought	Labeled as: <ul style="list-style-type: none"><li>• <i>rent</i> – failure to pay rent</li><li>• <i>breach</i> – violating a lease provision</li><li>• <i>illegal activity</i> – use of the property for an illegal purpose</li><li>• <i>at will termination</i> – for “tenants at will,”<sup>2</sup> who can be evicted for any reason or no reason, provided they were given proper notice</li><li>• <i>other</i> – for all other grounds</li></ul>

---

<sup>2</sup> This is a term for when a tenant does not have a written lease or stays after a written lease period with the consent of the landlord. A landlord can end a tenancy at will for any reason with 30-day’s notice.

Defenses	Whether the tenant asserted any defenses as a part of the proceeding	Labeled as: <ul style="list-style-type: none"> <li>• <i>bad process</i> – failure to follow proper procedures and timelines for the eviction</li> <li>• <i>inhabitable conditions</i> – a defense to a case based on failure to pay rent, used when the landlord let the unit fall into disrepair such that a tenant was excused for not paying some of the rent<sup>3</sup></li> <li>• <i>retaliatory eviction</i> – that the eviction is retaliation for protected actions<sup>4</sup></li> <li>• <i>discrimination</i> – that the landlord is using the proceeding as pretext for unlawful discrimination against the tenant</li> <li>• <i>no breach</i> – if the eviction is based on violating a lease provision, that the tenant did not violate the provision, or the provision is unenforceable under Mass. law</li> <li>• <i>other</i> – for all other grounds</li> </ul>
Landlord-tenant communications	Copies of any communication between landlords and tenants that were entered into the public record as part of a trial court case (obtained by scraping the case docket information for housing cases following the <a href="https://www.mass.gov/instructions-on-mass.gov">instructions on mass.gov</a> )	Text data of the docket available on <a href="https://www.masscourts.org">masscourts.org</a> , along with optical character recognition (OCR) extraction of the data within PDF files (any images or non-text data is discarded)

---

<sup>3</sup>In these cases, the housing court will order the landlord to make repairs, determine a fair value for the rent that is due in light of the poor conditions, and make the tenant pay only that value.

<sup>4</sup>This can include reporting a landlord to health officials, taking legal action against a landlord to enforce one’s rights, or organizing or joining a tenants’ rights organization.

Outcome	Whether the proceeding resulted in a court-ordered eviction or not.	<p>A binary “eviction” / “no eviction” input. Note that the company does not know the basis of the decision, and is aware that out-of-court settlements will be labeled as “no eviction,” even when the tenant leaves the property.</p> <p>At significant expense, the company could parse through a sample of the “no eviction” records to see what percentage are due to courts siding with tenants versus due to settlement, but would not know what the eventual outcomes in any settlement are.</p>
---------	---	--

### Database 2: Realtor Property Data

The second database has been created by the Lndlr.ly negotiating with most of the major realtors in the greater Boston area to create a custom database of rental property information, akin to the MLS databases that work with real estate transactions.

Field	Description	Value Type
Property location	Address, unit number, city, and zip code of the rental property	Stored as latitude/longitude, with an extra field for the unit number
Neighborhood	The name of the neighborhood commonly associated with the property location	Labeled based on the common names for the neighborhoods in the greater Boston area. (E.g. for Boston — Allston/Brighton, Back Bay, Dorchester, East Boston, Fenway/Kenmore, Hyde Park, Jamaica Plain, Roxbury, South End, etc.)
Bedrooms	The number of bedrooms in the unit	A number, with “0” indicating loft/studio apartments
Bathrooms	The number of bathrooms in the unit	A number, using the usual 1 for full baths, 0.5 for half-baths
Square footage	The square footage of the property	A number
Rent price	The listed rental price, with weighted modifiers to account for differences in included utilities across different properties	A number, in dollars

Turnover	The number of times the property has been listed as available in the past 10 years	A number
Amenities	Whether the unit includes any amenities, including a dedicated parking space, in-unit or in-building laundry, access to a fitness center, etc.	Open text field

### Database 3: Tenant Surveys

The third dataset is a series of customer surveys that LndLrd commissioned through market research firms. Tenants in the Boston area were recruited to participate through Facebook ads. Survey respondents were asked if they ever had an issue with their landlord that required them either leave a lease early or face an eviction proceeding. If they had, the respondent shared additional information.

Field	Description	Value Type
Landlord name	First and last name of the landlord, or entity name if the building is owned by a corporation.	Open text field
Property location	Address, unit number, city, and zip code of the rental property at issue	Stored as latitude/longitude, with an extra field for the unit number
Bedrooms	The number of bedrooms in the unit	A number, with "0" indicating loft/studio apartments
Bathrooms	The number of bathrooms in the unit	A number, using the usual 1 for full baths, 0.5 for half-baths
Rent price	The listed rental price, with weighted modifiers to account for differences in included utilities across different properties	A number, in dollars
Square footage	The square footage of the property	A number

Reason for dispute	The reason why the tenant breached the lease and left the property, or was evicted	Labeled as: <ul style="list-style-type: none"> <li>• <i>rent</i> – tenant could no longer pay rent</li> <li>• <i>tenant breach</i> – tenant violated a lease provision</li> <li>• <i>landlord breach</i> – landlord breached a lease provision</li> <li>• <i>illegal activity</i> – tenant used of the property for an illegal purpose</li> <li>• <i>bad conditions</i> – tenant found the conditions to be inhospitable</li> <li>• <i>personal incompatibility</i> – the landlord and tenant did not get along, to the point that the tenant had to leave</li> <li>• <i>change in circumstances</i> – tenant left for reasons that had nothing to do with the landlord, such as getting a job out of the Boston area</li> <li>• <i>Other</i> – for all other reasons</li> </ul>
Landlord-tenant communications	A copy of all communications (such as emails and public court documents) between the tenant and their landlord related to the dispute	Open text field and PDF attachments (any PDF attachments are then converted into text using optical character recognition, with images or any non-text data discarded)
Income level	Their income level, at the time the dispute arose	A number, in dollars
Education level	Their highest earned degree, at the time the dispute arose	Labeled as: <ul style="list-style-type: none"> <li>• <i>secondary education</i></li> <li>• <i>high school diploma or GED</i></li> <li>• <i>undergraduate</i></li> <li>• <i>masters</i></li> <li>• <i>professional degree</i> – including JD, MD, DMD</li> <li>• <i>PhD</i></li> <li>• <i>other</i></li> </ul>
Occupation	Their general field of work, based on the <a href="#">Standard Occupational Classification System</a> by the US Bureau of Labor Statistics	SOC number
Rental churn	The number of different units they had rented in the preceding 10 years	A number

Prior disputes	Did the tenant have a similar dispute with any prior landlords in the preceding 10 years?	A binary yes/no answer
Co-signor	Did the tenant have co-signor or other guarantor for the lease in question?	A binary yes/no answer