Beyond the Circle: Estimating Directional Commute Areas for On-Site Jobs using Bayesian Neural Networks

Atsushi Ide[1,*]

[1]SMS Co., Ltd., Tokyo, Japan

Abstract

When job seekers seek new opportunities, workplace location is often a primary concern. This is especially true in the nursing care sector, where the physical nature of assisting patients and elderly people necessitates on-site commuting. Generally, job recommender systems present postings within a circular area of arbitrary radius centered on the candidate's location. However, this circular approach fails to account for geographical nuances and the job seeker's specific directional preferences for their commute. To address this, we propose a novel method using a Bayesian Neural Network (BNN) to detect preferred commute areas by considering the job seeker's location, along with the direction and distance to potential facilities. On a large-scale evaluation of held-out job-seeker/facility pairs, our directional approach yields substantially more compact candidate areas than a fixed-radius baseline while capturing a higher density of relevant postings per unit area, at some cost to overall recall. An ablation against a deterministic Gradient Boosted Decision Tree baseline sharing the same direction-and-distance decomposition achieves comparable point-estimate accuracy, indicating that this efficiency gain stems primarily from explicitly modeling direction and distance rather than from the Bayesian machinery itself. We therefore argue that the practical value of the Bayesian formulation lies not in raw accuracy over deterministic alternatives, but in its calibrated uncertainty estimates, which support actionable, confidence-aware recommendations, such as encouraging job seekers to explore commonly overlooked but promising directions.

Keywords

Estimating commute areas, Bayesian Neural Networks, Direction aware, Uncertainty quantification, Human Resources

1. Introduction

While remote work is becoming increasingly prevalent, many industries still require a physical on-site presence. Specifically, in the nursing care sector where our services are mainly focused, nearly all job postings necessitate on-site work. Consequently, the commute area is a significantly more critical factor for these candidates than for those in other job categories. For this reason, we choose to model job seeker commute areas as a standalone preference, whereas many traditional recommender systems treat location merely as one of several simultaneous features, such as skills and experience.

Currently, a common approach for estimating a job seeker's commute area is to define a circular region with a fixed radius centered on their address. Candidate facilities are then selected from within this circle alongside other matching features. This method assumes that all facilities within the circle have an equal geographic probability of being selected. However, job seekers naturally possess specific directional preferences for their commute. For example, a candidate may avoid traveling west due to predictable traffic delays on a particular bridge or difficult driving conditions in mountainous terrain.

Since these directional biases vary depending on where a job seeker lives, it is more accurate to represent these preferences as a probability distribution. Furthermore, the acceptable commute distance within each specific direction is not absolute; it inherently contains a degree of uncertainty regarding length.

In this study, we propose a novel method to address these limitations by estimating a job seeker's preferred commute area through directional analysis. By applying a neural network to historical data encompassing user locations, the directions of favored facilities, and the distances, the network learns nuanced directional preferences. This not only allows for more accurate job matching but also provides a mechanism to mitigate geographic mismatch. By identifying a job seeker's true travel thresholds, recommender platforms can actively encourage candidates to explore viable facilities in previously overlooked directions, expanding their employment opportunities while optimizing the distribution of the workforce.

To support reproducibility, a repository containing our model implementation and evaluation scripts is available.[1]

2. Related Studies

The estimation of commute areas is critical in the nursing care sector, where on-site physical presence is mandatory. Traditional systems often rely on a circular area with an arbitrary radius centered on the job seeker. While early spatial models rely on isotropic distance decay functions to penalize far-away locations [1], they fail to capture the anisotropic (directional) reality of daily commutes. According to Activity Space Theory [2], person movement is affected by infrastructure and anchor points, supporting our observation that job seekers have distinct directional preferences due to obstacles like mountains or bridges. Trajectory mining studies [3] have similarly noted that human mobility is constrained by physical infrastructure rather than "as the crow flies" distance. Our directional Bayesian model, presented in the next section, offers a computationally efficient middle ground: it learns the latent geographic constraints of an area directly from user acceptance data, rather than calculating explicit road-network routing.

Furthermore, while early Location-Aware Recommender Systems (LARS) [4] introduced spatial indexing to improve efficiency, more recent advancements have integrated these spatial indices into deep learning architectures. For instance, recent studies have demonstrated the high efficiency of

RecSys in HR '26: The 6th Workshop on Recommender Systems for Human Resources, in conjunction with the 20th ACM Conference on Recommender Systems, September 28–October 2, 2026, Minneapolis, MN, United States.

*Corresponding author.
atsushi-ide@bm-sms.co.jp (A. Ide)
0009-0003-6999-2394 (A. Ide)

[1]Code and sample data available at: https://github.com/atsushi-ide1977/beyond-the-circle.git

utilizing H3 hexagonal coordinates [6] to generate location embeddings within Two-Tower recommendation models [7]. While these embedding-based approaches successfully demonstrate the efficacy of spatial data in improving retrieval scores, they typically project location into a dense latent space, treating geographic preference as an opaque similarity metric.

Our primary objective is fundamentally different. We aim to explicitly model and define the flexible geographic boundaries that reflect a user's directional commute preferences. Consequently, rather than converting spatial data into a latent embedding, our proposed method maintains an estimation of the commute area itself. By utilizing a Bayesian Neural Network [8, 9] to predict distinct directional probabilities and distance variances, we define interpretable geographic boundaries rather than relying solely on latent similarities. Ultimately, explicitly defining these directional boundaries not only improves recommendation accuracy but provides a tangible tool to counsel job seekers toward viable, under-explored regions, actively combating spatial mismatch.

3. Proposed Method

3.1. Preliminaries: Bayesian Uncertainty

As job search behavior is inherently stochastic, necessitating a model that can handle uncertainty, we build a model using the Bayesian Neural Network (BNN) framework with MC Dropout [10]. This allows the network to treat weights as probability distributions rather than fixed values. Following [5], we apply Heteroscedastic Aleatoric Uncertainty to model the variance in commute distances. During inference, we perform 200 stochastic forward passes to generate a robust predictive distribution. This ensures the system accounts for the noise in preferences, which is vital in job matching where geographic mismatch leads to unmatching or resigning at an early stage.

3.2. Learning Data

The raw dataset utilized for training comprises historical data collected between 2020 and 2025 in our service. This dataset consists of pairs of job seeker addresses and the corresponding locations of facilities where those individuals accepted an interview, serving as a definitive indicator of their real-world geographic constraints and preferences.

As illustrated in Figure 1, the process of preparing this learning data involves converting the spatial relationships between the job seekers and their accepted workplace facilities into structured features suitable for training the Bayesian Neural Network. The top part of the diagram shows how we utilize the coordinates of the job seeker's address, denoted as (lat_1, lon_1), and the facility address, (lat_2, lon_2), as raw inputs to derive our three primary features. These features are aggregated into the learning data table format shown at the bottom of Figure 1, using an example row ($lat_1 = 35.6758688$, $lon_1 = 139.7419168$, $direction = NE$, $distance = 9.5$) for clarity.

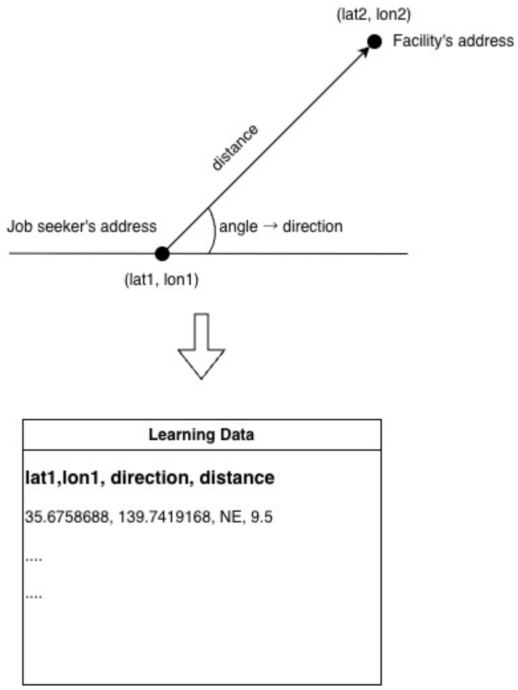

Figure 1: Schematic illustration of learning data.

- **User Location:** This feature consists of the raw coordinates of the job seeker's address (lat_1, lon_1), which serves as the reference origin point for the geometric calculations.
- **Direction (Categorical):** To capture directional preference, we first determine the geometric angle shown in Figure 1, calculated between the vector from the user's address toward the facility address and a true East reference vector. For model efficiency and interpretability, this continuous angle is discretized by classifying it into the closest of eight cardinal and ordinal directions (E, NE, N, NW, W, SW, S, SE). The example row demonstrates a derived direction of 'NE'.
- **Distance (Continuous):** This feature is the continuous distance from the user's address to the facility's address within each pair, calculated based on the geometric relationship between the two coordinates. As shown in the example row, this distance is recorded as a continuous value, such as 9.5 km.

Consequently, each training sample consists of the user location, the categorical direction, and the continuous commute distance. This structured $(lat_1, lon_1, direction, distance)$ approach, derived directly from the geometric relationships, allows the network to learn nuanced directional preferences.

3.3. Training Method and Prediction

3.3.1. Network Architecture

As illustrated in Figure 2, our model utilizes a multi-task learning architecture to simultaneously estimate directional preferences and commute distances. Given the input coordinates $x \in \mathbb{R}^2$, a shared neural network backbone extracts a latent spatial representation h. This shared representation branches into two specialized output heads:

1. **Direction Classification Head:** Outputs logits which are converted via a softmax function into

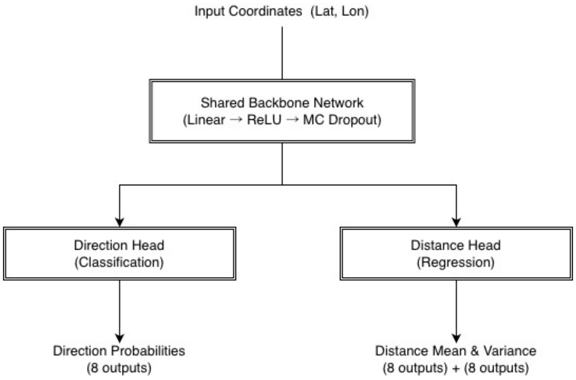

Figure 2: Multi-task learning architecture.

a probability distribution $p \in \mathbb{R}^8$ across the eight directional sectors.

2. **Distance Regression Head:** Outputs both the predicted mean distance $\mu \in \mathbb{R}^8$ and the log-variance $v = \log \sigma^2 \in \mathbb{R}^8$ for each of the eight directions to capture heteroscedastic aleatoric uncertainty.

3.3.2. Loss Formulation

To train the network, we optimize a multi-task loss function $\mathcal{L}_{total}$ that aggregates the classification loss $\mathcal{L}_{cls}$ and the regression loss $\mathcal{L}_{reg}$:

$$\mathcal{L}_{total} = \lambda_{cls}\mathcal{L}_{cls} + \lambda_{reg}\mathcal{L}_{reg} \tag{1}$$

For our experiments, we set both weighting hyperparameters λ_{cls} and λ_{reg} to 1.0, assigning equal importance to the learning of directional probabilities and distance variances, as this empirically yielded stable convergence.

The direction classification head is optimized using standard Cross-Entropy loss. Let p_j denote the predicted probability for direction $j \in \{1, \dots, 8\}$, obtained by applying a softmax activation to the raw logits output by the classification head. Given the ground-truth direction label y_{dir}, the classification loss is formulated as:

$$\mathcal{L}_{cls} = -\sum_{j=1}^{8} \mathbb{1}(y_{dir} = j) \log(p_j) \tag{2}$$

where $\mathbb{1}(\cdot)$ is the indicator function that evaluates to 1 when $y_{dir} = j$ and 0 otherwise. This objective penalizes the network when it assigns a low probability to the job seeker's actual accepted commute direction.

For the distance estimation, we apply an angular-proximity weighted Gaussian Negative Log-Likelihood (NLL). Because geographic space is continuous, a job seeker's distance tolerance in their true preferred direction k also provides implicit information about adjacent directions. Therefore, we supervise the distance head across all eight directions $j \in \{1, \dots, 8\}$ using a Gaussian weight w_j based on the angular difference $\Delta\theta_{k,j}$ between the true direction k and direction j:

$$w_j = \exp\left(-\frac{\Delta\theta_{k,j}^2}{2\sigma_{angle}^2}\right) \tag{3}$$

where σ_{angle} controls the sharpness of the angular falloff. The true commute distance y_{dist} is then evaluated against the predicted mean μ_j and predicted variance $\exp(v_j)$ for every direction, yielding the weighted regression loss:

$$\mathcal{L}_{reg} = \frac{\sum_{j=1}^{8} w_j \left(\frac{1}{2}\exp(-v_j)\,(y_{dist} - \mu_j)^2 + \frac{1}{2}v_j\right)}{\sum_{j=1}^{8} w_j} \tag{4}$$

This formulation allows the network to learn not only the mean expected distance but also the structural uncertainty of that distance. All networks were trained simultaneously, allowing the shared backbone to learn a robust geographic representation driven by both the categorical choice of direction and the continuous friction of distance.

4. Experiment

We evaluate the proposed method on a large-scale, held-out set of $n = 5{,}533$ job-seeker/facility pairs, each consisting of a job seeker's address and the location of a facility where that job seeker accepted an interview. These pairs were strictly withheld from the training dataset. The implementation of all three methods compared below, including the proposed BNN, the GBDT ablation, and the circular baseline, is available in the repository linked in Section 1.

For each pair, a target area was calculated using three methods:

1. **Proposed Method (BNN):** the Bayesian Neural Network described in Section 3, using 200 MC-Dropout stochastic forward passes at inference time.
2. **Deterministic Ablation (GBDT):** a Gradient Boosted Decision Tree baseline that receives the identical task decomposition as the BNN, namely direction classification and per-direction distance regression, but without MC Dropout or any explicit modeling of uncertainty. This ablation isolates how much of the improvement over the circular baseline is attributable to explicitly modeling direction and distance, as opposed to the Bayesian uncertainty-modeling machinery itself.
3. **Circular Baseline:** the area within a 7km radius circle centered on the job seeker's address, corresponding to the fixed radius currently used by our production commute filter.

We note that this historical data does not perfectly reflect job seekers' geographic preference alone. Candidate facilities were themselves selected through our traditional matching process, which primarily applies a fixed-radius circular filter and additionally allows human recommenders to introduce out-of-radius facilities that they judge to be a good match. Consequently, an accepted interview reflects not only distance and direction, but also this filtering and recommendation process, along with other factors such as job availability, posting exposure, salary, and schedule. We nonetheless use this data as a proxy for geographic preference because, regardless of these other factors, every job seeker must still consider the location of a facility before accepting an interview there.

Train/test splits were performed at the job-seeker level: each job seeker was assigned entirely to either the training set or the test set (an 80/20 split), so no job seeker appears in both partitions. The split is not, however, explicitly disjoint by facility or by time, so a given facility or time period may appear in both the training and test sets. We report this explicitly as a limitation on generalization claims (see Section 5.4).

To convert model outputs into a geographic area, we compute, for each of the eight directions, a destination point at the predicted mean distance along that direction's bearing from the job seeker's address. These eight destination points are connected in angular order to form a polygon, which is then filled with H3 cells at the target resolution; cells containing a destination point (and their immediate neighbors) are also included to guarantee edge coverage. All eight directions contribute to the polygon regardless of their predicted probability; the probability estimate is used only for downstream visualization and confidence signaling, not for excluding directions from the area.

All target areas were discretized into H3 hexagonal coordinates at resolution 9. To evaluate performance, we converted the true accepted job posting location into an H3 coordinate and checked whether it fell within the designated area for each method.

We report two complementary metrics. First, Spatial Hit Density is calculated by dividing the total number of correctly captured facilities by the total number of H3 cells generated by that respective method:

$$\text{Spatial Hit Density} = \frac{\text{Count of Captured Target Facilities}}{\text{Total Number of H3 Cells in Area}} \tag{5}$$

This metric serves as an index of recommendation efficiency, measuring the concentration of favorable facilities within the proposed area. A higher value indicates that the area contains a higher density of relevant jobs, reducing the noise or irrelevant postings presented to the user. However, because Spatial Hit Density rewards smaller areas and does not by itself reveal how many true accepted-interview pairs are actually captured, we report Recall as well, the fraction of all test pairs whose accepted facility falls inside the predicted area:

$$\text{Recall} = \frac{\text{Count of Captured Target Facilities}}{n} \tag{6}$$

Reporting both metrics together allows us to distinguish a method that is merely more *precise* (a smaller, denser area) from one that is also sufficiently *complete* (covering a large share of real commute outcomes).

5. Result

Table 1 summarizes the spatial evaluation across all $n = 5{,}533$ held-out pairs.

Both direction-aware methods, the proposed BNN and the deterministic GBDT ablation, substantially outperform the circular baseline in Spatial Hit Density (0.199 and 0.201, respectively, vs. 0.161) while using on average 33–35% fewer H3 cells per query (207.0 and 202.6 vs. 311.2). This confirms that explicitly modeling direction and distance, rather than assuming an isotropic circular boundary, produces a substantially more concentrated and efficient candidate search area.

This gain in density, however, comes with a trade-off in Recall: the circular baseline captures 50.0% of accepted-interview pairs simply by covering a larger area, while the BNN and GBDT capture 41.3% and 40.8%, respectively. We report this trade-off transparently rather than relying on Spatial Hit Density alone, since a system deployed in practice must balance the efficiency gains of a smaller search area against the risk of excluding a job seeker's true preferred facility. Nonetheless, both direction-aware methods each recovered a meaningful number of pairs that the circular baseline missed entirely (17 for the BNN and 23 for the GBDT ablation), i.e., candidates for whom the current production filter would have returned zero relevant postings at all. This indicates that, even where overall recall is lower, the directional models can provide qualitatively different, and in these cases strictly better, recommendations than the circular filter.

Table 1
Spatial evaluation results across $n = 5{,}533$ held-out pairs. "Avg. Cells" is the mean number of H3 cells generated per query; "Density" is Spatial Hit Density (Eq. 5), reported as a raw ratio (hits per H3 cell), not a percentage; "Recall" (Eq. 6) is reported as a percentage.

Method	Avg. Cells	Hits	Density	Recall
Proposed (BNN)	207.0	2283	0.199	41.3%
Ablation (GBDT)	202.6	2257	0.201	40.8%
Circular (7km)	311.2	2766	0.161	50.0%

5.1. Ablation Study: Isolating the Value of Bayesian Modeling

To address whether the improvement over the circular baseline stems specifically from the Bayesian uncertainty-modeling apparatus (MC Dropout, heteroscedastic aleatoric variance), or simply from the underlying task decomposition into direction and distance, we compare the proposed BNN against the deterministic GBDT ablation described in Section 4. The two methods perform comparably: the GBDT ablation attains a marginally higher Spatial Hit Density (0.201 vs. 0.199) and a smaller average area (202.6 vs. 207.0 cells), while the BNN captures slightly more absolute hits (2283 vs. 2257, a difference of only 26 pairs, or 0.5% of n) and recovers fewer circular-baseline misses than the GBDT (17 vs. 23). Given the small magnitude of these differences relative to the sample size, we do not interpret either method as more accurate than the other in this evaluation.

This result is informative rather than discouraging: it indicates that the primary source of improvement over the circular baseline is the explicit direction-and-distance decomposition itself, which both models share, rather than the Bayesian machinery in isolation. We therefore do not claim the BNN is more accurate than a well-tuned deterministic alternative. Instead, we position the value of the Bayesian formulation along a different, complementary axis: unlike the deterministic GBDT, the BNN additionally yields calibrated epistemic and aleatoric uncertainty estimates for every direction. In a deployed recommender system, this uncertainty enables concrete downstream behaviors that raw point estimates cannot support, such as withholding or flagging low-confidence directional boundaries, communicating a distance range rather than a single number to recruiters or job seekers, and prioritizing directions in which the model is both accurate and confident. We view this interpretability and uncertainty quantification, rather than a raw accuracy gain over deterministic alternatives, as the primary practical justification for the added complexity of the Bayesian approach. We have not yet directly validated this claim through a calibration study (e.g., reliability diagrams) or a downstream uncertainty-aware experiment; we treat this as an important next step for substantiating the practical value of the Bayesian formulation.

5.2. Error Analysis

We examined pairs that yielded differing results, paying special attention to instances where the proposed method failed to cover the target facility while the traditional methods succeeded. In the majority of these divergent cases, the area generated by the proposed method was located adjacent to, or very close to, the H3 cell containing the accepted facility. This suggests an edge-effect inherent to discrete global grid systems; while the model accurately predicted the general vector and distance, the rigid boundaries of the H3 cell slightly missed the specific coordinate.

5.3. Qualitative Visualization

To illustrate the flexibility of the proposed method compared to rigid geometric baselines, Figure 3 visualizes a sample prediction for a job seeker, overlaid on an H3 hexagonal grid (resolution 8). The plotted nodes represent the model's predicted mean commute distance for each of the eight discrete directions. Nodes marked in grey indicate directions that the classification head determined have a low probability of acceptance. In this specific example, these low-probability vectors directly correlate with real-world geographical constraints: a mountainous region to the west and an area with inconvenient traffic routing to the north. By connecting the high-probability distance boundaries, the model generates an asymmetric, irregular polygon that accurately reflects the user's commute preferences. The H3 cells falling within this boundary are highlighted in yellow to represent the final target area. Furthermore, the varying color density among these highlighted cells visually encodes probabilities from the model; cells with denser coloring represent a relatively higher predicted probability of the job seeker accepting a commute in that specific direction. This granular, probability-weighted area allows the recommendation system to prioritize facilities in the darkest regions to lead to better matchings.

5.4. Limitations and Future Work

Our approach has several limitations that we leave for future work. First, although our training data spans locations across Japan nationwide, sample density is uneven across the country: some rural areas are represented by comparatively few historical acceptance pairs, and the model's directional and distance estimates in these lower-density areas are correspondingly less reliable than in densely sampled urban regions. More broadly, the model learns direction and distance preferences from historical acceptance patterns tied to Japan's road network, geography, and transit infrastructure, so it is not obvious how to transfer it to a country or region with substantially different infrastructure without local training data. A promising solution is to treat the trained network as a source model and apply transfer learning or fine-tuning with a small amount of local acceptance data from an under-sampled or new region, rather than training from scratch, or to incorporate region-level

Figure 3: Spatial visualization of a predicted commute area for a sample job seeker.

features (e.g., population density, public transit density, road network typology) that could allow the shared backbone to generalize more robustly across regions with sparse data.

Second, the desirability of a given route or direction may drift over time, for example due to new infrastructure, road closures. Our current model is trained on a static historical snapshot (2020–2025) and does not explicitly model this temporal non-stationarity. Periodic retraining, or an explicit time-decay weighting of older acceptance pairs, would help the model remain calibrated to current commute conditions.

Third, our model conditions only on the job seeker's residential location, and therefore learns *location-conditioned aggregate commute behavior* rather than a truly individualized preference model. Job seekers who share a residential area are treated as having identical directional and distance preferences, even though individual factors (e.g., car ownership, personal mobility constraints) may cause their true preferences to differ from the local aggregate. Extending the model with individual-level features, where available, is a natural next step toward genuine personalization.

Fourth, our train/test split (Section 4) is disjoint by job seeker but not by facility or time, and our evaluation reports Spatial Hit Density and Recall at a single operating point per method rather than recall–area trade-off curves, matched-operating-point comparisons, or statistical significance tests. We view Density as an efficiency indicator rather than a precision-equivalent metric, and leave a more rigorous evaluation protocol, including leakage-safe splits and stronger baselines, to future work.

Finally, our distance feature is currently derived from geodesic (straight-line) distance rather than actual travel time or route distance. Incorporating a road-network travel-time estimate, e.g., via a routing API or an offline network such as OpenStreetMap, as an additional input feature is a natural extension that could sharpen the distance-regression head, particularly in areas where geodesic distance is a poor proxy for commute effort (e.g., rivers, mountains, limited bridge crossings).

6. Conclusion

We demonstrated that our proposed method, which models commute areas based on directional preferences and distances, is effective for estimating more compact and efficient geographic boundaries for job seekers than a fixed-radius circle, learning location-conditioned aggregate commute behavior directly from historical acceptance data. As illustrated by our probability-weighted spatial mapping, visualizing these preferable areas with directional granularity provides a dual benefit. First, it improves the efficiency of recommending facilities that job seekers from a given area are historically inclined to accept. Second, it offers interpretable, actionable insights for career counseling. By visually presenting commute thresholds typical of job seekers from their area, and revealing how physical constraints shape historical job search behavior, advisors or recommender platforms can encourage job seekers to consider viable opportunities in less popular or overlooked directions, thereby expanding their options and increasing their chances of successful placement. Ultimately, this approach could serve our broader societal mission: mitigating the geographic maldistribution of essential workers. Systematically guiding candidates toward viable roles outside of highly concentrated urban centers could help to ensure that vital nursing care reaches underserved rural communities.

Declaration on Generative AI

Gemini 3.1 was used to identify and correct grammatical errors, typos, and other writing mistakes, as well as to refine word choice while drafting this paper. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] Mao Ye, Peifeng Yin, and Wang-Chien Lee. 2010. Location recommendation for location-based social networks. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 458–461. https://doi.org/10.1145/1869790.1869806

[2] Stefan Schönfelder and Kay W. Axhausen. 2003. Activity spaces: Measures of social exclusion? *Transport Policy*, Elsevier, vol. 10(4). 273–286.

[3] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining interesting locations and travel sequences from GPS trajectories. In *WWW '09: Proceedings of the 18th International Conference on World Wide Web*. 791–800. https://doi.org/10.1145/1526709.1526816

[4] J. J. Levandoski, M. Sarwat, A. Eldawy, and M. F. Mokbel. 2012. LARS: A Location-Aware Recommender System. In *IEEE 28th International Conference on Data Engineering*, Arlington. 450–461. https://doi.org/10.1109/ICDE.2012.54

[5] Alex Kendall and Yarin Gal. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? *arXiv:1703.04977*. https://arxiv.org/abs/1703.04977

[6] Uber. 2023. H3: Hexagonal hierarchical geospatial indexing system. https://github.com/uber/h3

[7] Ivo Silva, Guilherme Bonaldo, and Pedro Nogueira. 2025. Location Matters: Leveraging Multi-Resolution Geo-Embeddings for Housing Search. In *Proceedings of the Nineteenth ACM Conference on Recommender Systems*. 1006–1009. https://dl.acm.org/doi/epdf/10.1145/3705328

[8] Alex Graves. 2011. Practical variational inference for neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems*. 2348–2356. https://dl.acm.org/doi/10.5555/2986459.2986721

[9] Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. 2015. Weight uncertainty in neural network. In *Proceedings of the 32nd International Conference on Machine Learning*. 1613–1622. https://dl.acm.org/doi/10.5555/3045118.3045290

[10] Yarin Gal and Zoubin Ghahramani. 2016. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *arXiv:1506.02142*. https://arxiv.org/pdf/1506.02142