UDC 338.51

Kaiyrbek Zhanadil
Grade 11 student
Scientific supervisor: Nurgaliyeva Almagul Adilbekovna
Nazarbayev Intellectual School of
Science and Mathematics in Nura district of Astana
(Astana, Kazakhstan)

HOW ACCURATELY CAN OBSERVABLE LISTING DATA ESTIMATE FAIR PRICES FOR USED CARS IN KAZAKHSTAN? A HEDONIC MODEL OF 2,546 ONLINE LISTINGS

Abstract: A buyer who sees a used car priced far below similar cars has to know first what similar cars cost. This study asks how accurately a model that uses only what a listing shows can estimate the asking price of a used car in Kazakhstan. It uses a dataset of kolesa.kz listings compiled and published online by a third party (version of 18 December 2025): after duplicates, new cars and listings removed by cleaning rules written before the data were obtained (refined on 12 October after the matched texts were read), 2,546 used-car listings remain. A hedonic ridge regression of the logarithm of the asking price on brand and model, age, mileage, engine volume, gearbox, fuel, body type and steering side was compared with a simple benchmark, the median price of the same brand, model and year, using five-fold cross-validation. The model explained 86.2% of the variation in log price (benchmark 60.3%), and its typical (median) error was 16.7% of the asking price (benchmark 22.5%); 57.1% of the estimates were within 20% of the asking price. Where listings of the same brand, model and year existed (1,386), their median was as precise as the model (12.9% vs 13.7%); the model's advantage comes from cars without such comparables. Errors were largest for old cars: the median error rose from 12.6% for cars up to 3 years old to 29.9% for cars over 30 years old. Once the other fields were known, removing age cost the most accuracy (R^2 down by 0.175); each additional year of age was associated with a 6.8% lower asking price, close to the 7% reported for Kazakhstan in 2020, while mileage added almost nothing once age was known. Of all listings, 18.9% were priced at least 20% below the estimate and 10.8% at least 30% below, and gaps of the same size on a ratio scale occur about equally often above the estimate, so a price far below the estimate does not by itself show a bargain or a hidden defect.

Keywords: used cars, asking prices, hedonic regression, price estimation, cross-validation, online listings, Kazakhstan

Introduction

Buyers of used cars know less about a car than its seller does. Akerlof showed, with the used-car market as his main example, that when buyers cannot tell good cars from bad ones ("lemons"), they offer an average price, and owners of good cars may leave the market [1]. Online listing platforms change this situation in two ways. Sellers can disclose much more about a car, and on eBay Motors such disclosures were found to be important determinants of the price [2]. Platforms also put thousands of listings side by side, so a buyer can sort them by price and compare a car with similar ones.

This study is motivated by that comparison. In a hypothesis written before any data were obtained, the author proposed three reasons why a used car may be listed far below similar cars: the seller needs money quickly, the seller openly admits a defect, or the seller hides a defect and explains the low price by urgency. Before any of these reasons can be studied, a more basic question has to be answered: what is the price of "similar cars" that a listing is compared with, and how precisely can it be estimated from what a listing shows?

Predicting used-car prices from listing data is a well-studied task. A recent study comparing machine-learning models on a dataset of 4,009 used cars reported that the R^2 of a stacking regressor rose from 0.14 on raw data to 0.8899 after careful feature engineering [3]. For Kazakhstan, Barlybayev and co-authors collected 178,397 listings from kolesa.kz on 13 March 2023 with 16 attributes and compared twenty models; bagged trees reached a correlation of $R = 0.94$ between predicted and actual prices, while a plain linear model reached $R = 0.48$ [4]. Another study reported $R^2 = 0.9542$ for a random forest and 0.8918 for linear regression, but its data were 200 records generated from realistic Indian market price ranges, not real listings [5]. These three studies, found by the author, show that reported accuracy depends strongly on the data and on the measure used, so numbers from different studies cannot be compared directly. Hybrid models that combine gradient boosting, random forests and linear regression have also been proposed [6], and a review notes that tree-ensemble models such as XGBoost and random forests struggle with luxury vehicles, whose price depends on brand reputation, rarity and configuration, and suggests treating price tiers separately [7].

The closest work on Kazakhstan's market is a 2020 study by Bekishev published by AERC, which estimated a hedonic regression on 57,205 internet listings with 13 characteristics. It found, for example, that each year of age lowers the price by 7%, each additional 100,000 km by 2%, right-hand drive by 10% and the absence of customs clearance in Kazakhstan by 33%, while an automatic gearbox raises it by 28% [8]. The hedonic approach explains the price of a product by the implicit prices of its characteristics [9].

These studies either report how prices depend on characteristics [8] or report the accuracy of a model in measures such as R or RMSE [4]. A buyer, however, needs to know how far a typical estimate is from the asking price, whether a simple rule such as "the median for the same model and year" does almost as well, for which cars the estimate is least reliable, and which fields of the listing carry the information. This study answers these points on a dataset published in December 2025. Its research question is: how accurately can a model using only observable listing data estimate fair prices for used cars in Kazakhstan? In this paper the "fair price" of a listing is the asking price that the model expects for a car with the same visible characteristics. It is a typical asking price for such a car, not a valuation of the specific car and not a transaction price.

Materials and Methods

Data. The study uses the dataset "Kolesa-Cars-2025", compiled from kolesa.kz listings and published online by the user mutashevdarkhan (version 1, last updated 18 December 2025) [10]; the MIT licence shown on the dataset page is the label chosen by the uploader. Its single file, cars_kolesa_parsed.csv, has 3,999 rows and 14 columns: the listing title, the asking price, the price as displayed, a short text from the listing, the link to the listing on kolesa.kz, brand, model, condition (new or used), model year, mileage, engine volume, fuel, gearbox and drive. The dataset page gives no description of how or when the listings were collected. All

prices are asking prices in tenge. The file has no city, no listing date and no customs status, and the drive column is empty in all but 35 rows. The text column holds only the beginning of each listing, apparently the text shown in search results: its median length is 121 characters, and 86.2% of the texts in the final sample are cut off. No listing data were collected automatically by the author; apart from the manual check of current listings described in the Discussion, all data come from [10].

Sample. Table 1 shows how the sample was built. Several rows can describe the same listing, so one row was kept for each listing number in the link. New cars were removed. Then the cleaning rules written on 28 September 2026, before any data had been obtained, were applied in the stated order to the text visible in each listing: a price that is not a full sale price («первый взнос», «в кредит»), an exchange offer («обмен»), a car sold for parts or not running («на запчасти», «не на ходу»), a repost of the same car (defined here as the same brand, model, year, mileage, engine volume and price under another listing number; listings without mileage were not checked; the one with the lower listing number was kept), and an implausible price (below 10% of the median price of the same brand and model). Three details were added on 12 October 2026, after the matched texts were read: the keywords «первоначальн» and «в рассрочку» were added to the first rule, and «в кредит» was counted only as a way of paying, not as «в кредите», which says that the car is under a loan; a keyword was not counted when it was negated (for example «обмена нет», «без обмена», «обмен жок», «обмен не интересует»); and «досталась обменом» (the car was obtained by exchange) was not counted as an exchange offer. Every matched text was read and classified by the author's research mentor with an AI assistant. Misspelled or Kazakh-language forms such as «не нахаду» and «жүрмей тұр» are not caught by these rules. The exchange rule removed 35 listings that offer an exchange, including listings where the price is for cash and an exchange costs more, and kept 12 in which an exchange is refused or only mentioned. The file does not record the generation of a model, so a model's median can mix several generations; the last rule removed a 1995 Audi A6 listed at 10,000 tenge and a 2002 Mercedes-Benz S 600 listed at 5.3 million tenge, whose model median (113.65 million tenge) is the midpoint of only two listings, the other a 1993 S 600 listed at 222 million tenge, so the rule removed a plausible price and kept the extreme one.

Table 1 - Construction of the sample

Step	Listings left
Rows in the file	3,999
One row per listing number (968 duplicate rows removed)	3,031
Used cars only (444 new cars removed)	2,587
Not a full sale price (1 removed)	2,586
Exchange offer (35 removed)	2,551
For parts or not running (1 removed)	2,550
Repost of the same car (2 removed)	2,548
Implausible price (2 removed)	2,546

Features. Every feature is something a buyer sees in the listing: brand; brand and model together (models with fewer than three listings were pooled as "other models" of their brand, which gives 256 model groups); age, defined as 2025 minus the model year, and its square; the logarithm of mileage with a flag for missing mileage; engine volume; fuel; gearbox; body

type; right-hand drive; and whether the car is offered "to order" (на заказ). Engine volume was taken from its column or, when the column was empty, from the "N л" item of the listing text; in the 1,851 listings where both are present they agree in every case. Body type, right-hand drive and "to order" were read from the fixed opening items of the listing text.

Estimates compared. Three estimates of each listing's price were computed. The first (B0) is the median price of all listings, a reference that uses no information about the car. The second (B1) is a simple benchmark close to what a buyer would do by hand: the median price of listings of the same brand, model and model year; when no such listing exists, the median for the same brand and model, then for the same brand, then for all listings. The third (H) is a hedonic regression of the logarithm of the price on all the features above, estimated as a ridge regression, which shrinks the coefficients towards zero and keeps the estimates for rare models from being driven by one or two listings [11, p. 237]. The penalty was set at 1, the value used in the preliminary check of 11 October 2026 (numeric features standardised), and the values 0.1 and 10 are reported as a check. A simple baseline was compared with a more advanced model, as a recent study in this journal also did [12], where logistic regression served as "the primary baseline" for more complex fraud-detection models.

Validation. All three estimates were evaluated with five-fold cross-validation [11, p. 203]: the listings were split at random (seed 20261012) into five parts of nearly equal size, and each part was priced by an estimate built only on the other four. Every listing therefore receives an out-of-fold estimate, made from data that did not include it. Accuracy is reported as the R^2 of the log price, the median absolute percentage error (the typical distance between the estimate and the asking price, as a share of the asking price), the mean absolute percentage error and the share of listings estimated within 10% and within 20%. Ninety-five percent confidence intervals were obtained by resampling listings 1,000 times (the models were not refitted within resamples). To check stability, the five-fold split was repeated with ten other seeds; to check the cleaning rules, the model was also run on all 2,587 used-car listings. The preliminary check on the same file on 11 October 2026 had already shown that a model of this kind reaches an R^2 of about 0.85, so this analysis was not blind to the data.

Which fields matter. Three simple checks were made. First, each group of fields was used on its own. Second, groups were added to the model one at a time in the order in which the author ranked them on 2 October 2026, before the data were obtained: brand and model, then year (he wrote "generation and year"; the file has only the year), mileage, engine volume, gearbox and steering side (he also listed drive, which is almost always empty in the file, and customs status, which is absent), and finally the fields he did not list (fuel, body type, "to order"). Third, each group was removed from the full model in turn. Because many fields move together (older cars have higher mileage), the third check measures what a field adds when all the others are already known, not its total importance.

Comparison with earlier estimates. To compare with Bekishev's percentages [8], the same features were used in an ordinary least-squares regression in which age and mileage (per 100,000 km) enter linearly and every brand-model group has its own level. The percent effect of a feature is exp(b) - 1, with a 95% interval from 500 resamples.

Listings far from the estimate. As a descriptive result only, the listings priced at least 20%, 30% and 40% below the estimate of H were counted, together with the listings as far above it on a ratio scale, that is, with an asking price at least 1/0.8 = 1.25, 1/0.7 (about 1.43) and 1/0.6 (about 1.67) times the estimate; the plain counts of listings at least 20% and 30%

above the estimate are also given, and the counts below were repeated for B1. The study does not claim that such listings are bargains, errors or hidden defects.

Change from the original design. The design written on 28 September 2026, before data were available, named a different main method: the price of a listing would be the median of at least five comparable listings of the same brand, model and generation, model year within one year, mileage within 30,000 km and the same engine volume, gearbox, drive, fuel and city. The author drafted this rule (sent on 27 September 2026: brand and model, model year within one year, mileage within 15,000 km, engine volume and drive, customs status), and his mentor revised it on 27 and 28 September (mileage window 30,000 km; gearbox, fuel and city added; customs status left out, because it is one of the possible reasons for a low price). Because this file records no city and no generation and almost no drive, the rule was applied without these three conditions, and the number of listings it can price is reported below. Since it can price only a small share of the listings, the hedonic model became the main method. The design also planned to restrict the study to the three to five most listed models in one or two cities; because the file has no city and only a few thousand listings, all used cars were kept. The design's main threshold of 30% below the expected price is reported together with its 20% and 40% thresholds. All computations were done in Python 3.13 with pandas 3.0.3 and NumPy 2.5.1; the figures were drawn with Matplotlib 3.11.0.

Results

Sample. The 2,546 listings cover 67 brands and 526 brand-model pairs. The median asking price is 6.5 million tenge (interquartile range 3.5 to 11.5 million), and model years range from 1956 to 2025 (median 2013). The most frequent brands are Toyota (560 listings), Hyundai (289), Mercedes-Benz (183), VAZ (Lada) (175), Kia (173) and BMW (146). Automatic gearboxes account for 1,730 listings and manual ones for 608; 125 cars have right-hand drive. Mileage is missing in 485 listings; engine volume was filled from the listing text in 680 and is still missing in 15.

Accuracy. Table 2 compares the three estimates. The hedonic model explained 86.2% of the variation in log price and had a median error of 16.7% of the asking price. The benchmark of the same brand, model and year had a median error of 22.5% and explained 60.3%. The difference in median error, 5.8 percentage points, has a 95% interval of 4.5 to 7.0 points, and the hedonic estimate was closer to the asking price than the benchmark for 61.9% of the listings. The benchmark could use the same brand, model and year for 1,386 listings; for the rest it fell back to the brand and model (865), the brand (277) or all listings (18). On the 1,386 listings with comparables of the same brand, model and year, the benchmark was as precise as the model (median error 12.9% against 13.7%; R^2 0.876 against 0.910), while on the fallback listings it was far worse: 38.5% against 18.9% at the brand-model level, 50.0% against 29.1% at the brand level and 55.7% against 38.2% for the 18 priced by the median of all listings. The model's advantage therefore comes from the 1,160 cars without such comparables.

Table 2 - Out-of-fold accuracy of three estimates of the asking price (n = 2,546)

Estimate	R^2 of log price (95% CI)	Median error, % (95% CI)	Within 10%, %	Within 20%, %
B0: median of all listings	-0.003	51.9 (50.0 to 54.8)	8.4	18.6
B1: median of the same brand, model and year	0.603 (0.566 to 0.638)	22.5 (21.2 to 23.7)	27.3	46.7
H: hedonic ridge regression	0.862 (0.837 to 0.881)	16.7 (15.9 to 17.6)	32.7	57.1

The result is stable. Over ten other random splits, the R^2 of the hedonic model ranged from 0.860 to 0.867 and its median error from 16.3% to 17.0%. With a penalty of 0.1 or 10 the R^2 was 0.855 and 0.858 and the median error 16.3% and 17.4%. Without any cleaning rules, on all 2,587 used-car listings, the R^2 was 0.854 and the median error 16.8%. The mean error of the hedonic model, 25.0%, is much larger than its median error, because a small number of listings are very far from their estimates (Figure 1).

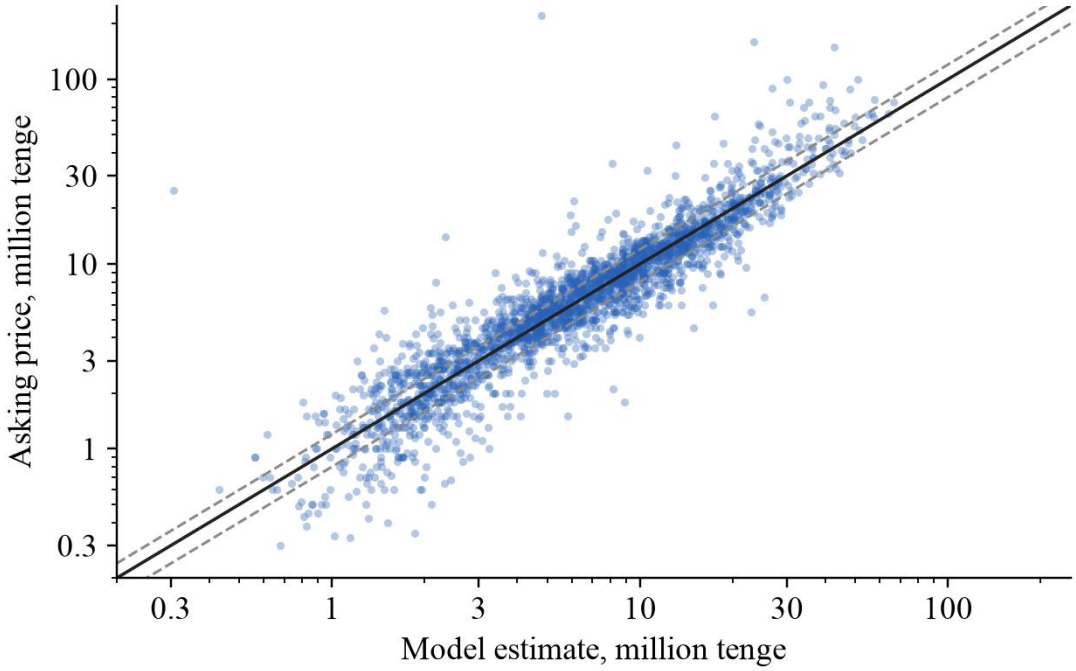


Figure 1 - Asking price and out-of-fold estimate of the hedonic model for 2,546 listings, million tenge, logarithmic scales. Solid line: estimate equal to the asking price; dashed lines: asking price 20% above and 20% below the estimate

Where the model errors. Table 3 shows the error by price band. The estimate was most precise for cars listed between 10 and 20 million tenge (median error 12.9%) and least precise for cars under 2 million tenge (29.6%). For the cheapest cars the estimate was usually too high (median estimate 23.5% above the asking price), and for cars of 20 million tenge and more it was usually too low (11.6% below). Almost all of this pattern is mechanical, because grouping listings by their own asking price puts unusually cheap listings in the lowest band and unusually expensive ones in the highest: grouped by the estimate, the median signed error is +0.3% under 2 million and -0.9% at 20 million and over (Table 3). By brand (brands with at least 50 listings), the median error ranged from 9.8% for Hyundai and 11.4% for Lexus to

24.2% for BMW, 24.9% for Audi, 25.9% for GAZ and 26.9% for Mercedes-Benz. The high errors for Audi and Mercedes-Benz do not come from luxury cars: the median Audi in the sample is 31 years old and listed at 2.2 million tenge, and the median Mercedes-Benz is 23 years old and listed at 5.5 million tenge. Across all listings, the error grows with the age of the car: the median error is 12.6% for cars up to 3 years old (402 listings), 13.2% for 4 to 10 years (608), 17.2% for 11 to 20 years (918), 21.1% for 21 to 30 years (429) and 29.9% for cars over 30 years old (189). The listings furthest above their estimates include collector cars, such as a 1956 GAZ M-20 Pobeda listed at about 25.0 million tenge with an estimate of 0.31 million, and recent luxury cars.

Table 3 - Error by price band (listings and errors grouped by the asking price; signed error: median of the estimate relative to the asking price, positive when the estimate is higher, grouped by the asking price or by the estimate)

Price band, million tenge	Listings	Median error, H, %	Median error, B1, %	Signed error, H, by asking price, %	Signed error, H, by estimate, % (listings)
under 2	308	29.6	53.3	23.5	0.3 (309)
2 to 5	655	18.2	27.1	-1.6	-4.4 (701)
5 to 10	810	15.1	17.2	-3.6	-1.6 (761)
10 to 20	543	12.9	17.4	-0.8	3.5 (554)
20 and over	230	17.7	26.8	-11.6	-0.9 (221)

Which fields matter. Table 4 shows the three checks. On its own, brand and model explained 53.6% of the variation in log price, with a median error of 32.2%, and age explained 44.8%. Adding the year to brand and model raised R^2 to 0.836 and cut the median error to 17.1%. Every later group, in the author's order, changed R^2 by 0.013 or less and the median error by less than half a percentage point. When groups were removed from the full model, age caused by far the largest loss (R^2 down by 0.175, median error up by 11.7 points), followed by brand and model (0.073 and 7.3 points). Removing mileage, steering side, fuel or "to order" changed R^2 by 0.0015 or less.

Table 4 - Contribution of field groups to the hedonic model (out-of-fold)

Field group	R^2 alone	R^2 after adding (author's order)	Median error after adding, %	R^2 lost when removed	Median error added when removed, points
Brand and model	0.536	0.536	32.2	0.073	7.3
Year (age)	0.448	0.836	17.1	0.175	11.7
Mileage	0.173	0.837	17.0	0.001	0.0
Engine volume	0.204	0.850	17.3	0.006	0.0
Gearbox	0.303	0.854	16.9	0.007	0.0
Steering side	0.021	0.855	17.0	0.000	-0.1
Fuel, body type, "to order"	0.015 / 0.229 / 0.003	0.862	16.7	0.000 / 0.007 / 0.000	-0.1 / 0.5 / -0.1

Comparison with Bekishev. Table 5 sets the implied effects of this dataset next to those reported for 2020 [8]. The effect of age is almost the same: 6.8% per year here and 7% there. An automatic gearbox is associated with a 21.3% higher price, somewhat less than 28%. A higher mileage lowers the price more here (5.6% per 100,000 km against 2%). The effect of right-hand drive, -3.8%, is uncertain, because only 125 listings have it, and its interval includes zero. The largest difference is in engine volume: 1.1% per 0.1 litre here against 21% in the 2020 study.

Table 5 - Percent effects on the asking price in this study and in Bekishev (2020)

Feature	This study, % (95% CI)	Bekishev 2020 [8], %
One more year of age	-6.8 (-7.2 to -6.4)	-7
100,000 km more mileage	-5.6 (-7.6 to -3.8)	-2
Automatic instead of manual gearbox	+21.3 (+14.4 to +29.2)	+28
Right-hand drive	-3.8 (-14.7 to +5.2)	-10
Engine volume 0.1 litre larger	+1.1 (+0.5 to +1.9)	+21
Not cleared through customs in Kazakhstan	not in the data	-33

Listings far from the estimate. Of the 2,546 listings, 480 (18.9%, 95% CI 17.4 to 20.4) were priced at least 20% below the hedonic estimate, 276 (10.8%, 95% CI 9.6 to 12.1) at least 30% below and 151 (5.9%) at least 40% below (Figure 2). On a ratio scale, 20% below the estimate is as far from it as 25% above, 30% below as about 43% above, and 40% below as about 67% above. That far above the estimate were 515 listings (20.2%), 253 (9.9%) and 123 (4.8%), so gaps of this size occur about equally often in both directions. In plain percentages, 621 listings (24.4%) were priced at least 20% above the estimate and 415 (16.3%) at least 30% above. Against the simpler benchmark B1 more listings are far below: 652 (25.6%) at least 20% below and 476 (18.7%) at least 30% below. Only 184 listings are at least 30% below under both estimates.

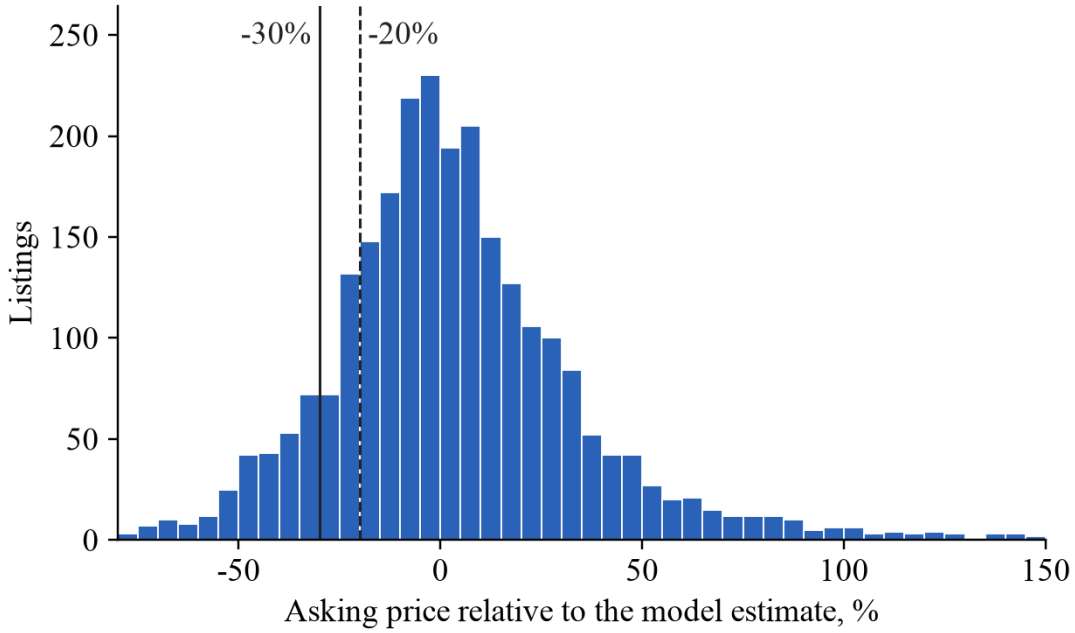


Figure 2 - Asking price relative to the out-of-fold estimate of the hedonic model, % (2,522 listings between -80% and +150% are shown; 23 listings above +150% and 1 below -80% are outside the range). Lines mark 20% and 30% below the estimate

The comparable-car method of the original design. Of the 2,047 listings with mileage, engine volume and fuel recorded, 138 (5.4% of the sample) had at least five comparable listings under the design's rule without city, generation and drive. On these 138 listings the median of comparable listings had a median error of 8.2%, and the hedonic model 8.6%. The comparable-car method is therefore precise for popular, well-described cars, but it cannot price about 95% of this sample.

Discussion

The answer to the research question is that, in this dataset, the visible fields of a listing estimate the asking price of a used car with a typical error of about 17%. Where listings of the same brand, model and year existed (1,386), their median was as precise as the model (12.9% vs 13.7%); the model's advantage comes from cars without such comparables, which make up 1,160 of the 2,546 listings, so for a car with listings of the same model and year the model adds little, while for other cars it is clearly better. This is the precision of a typical estimate; for a single listing the error can be far larger, as the gap between the median error (16.7%) and the mean error (25.0%) shows. The estimate is least reliable for old cars: the median error more than doubles from cars up to 3 years old (12.6%) to cars over 30 years old (29.9%), which may reflect differences in condition that the listing fields do not show.

The ranking of fields agrees in part with the ranking the author gave before seeing the data. He put brand and model first and the year second. On their own, brand and model explain more of the variation (53.6%) than age (44.8%), and together they explain 83.6%. He put mileage third, but once the age of a car is known, mileage adds almost nothing to the accuracy. One reason may be that mileage follows age closely; another is that mileage is missing in 485 listings, and 18 listings of cars at least three years old report less than 1,000 km, which is not plausible (treating these 18 as missing changed R^2 from 0.862 to 0.861). The

regression still gives mileage a clear effect of 5.6% per 100,000 km, but this information is mostly already contained in age.

The close agreement with Bekishev on age (6.8% against 7% per year) is notable, because his study was published almost six years before this dataset [8]. The other differences may come from the different models. Bekishev's list of 13 characteristics includes the brand but not the model [8], whereas here every brand-model group has its own price level, so engine volume is compared only within the same model. Within a model, a larger engine changes the price only a little; across models, a larger engine also stands for a larger or more expensive car, which may explain why his estimate for engine volume (21% per 0.1 litre) is so much larger than the one found here.

The author's hypothesis, written on 24 September 2026 before any data were obtained, was that cars listed below the market value may be priced below it for three reasons: a forced sale within a short time, a defect that the seller openly admits, or a defect that the seller hides while explaining the low price by urgency. This dataset cannot test these reasons, because none of them is a field of the listing. What it can show is their possible scale. The visible fields explain about 86% of the variation in log asking prices, so about 14% is left for everything else: condition, accident history, the city, customs status, the seller's urgency, plain differences in how sellers set prices, and the imperfections of the model itself. It also shows that a price far below the estimate is common: about one listing in ten is at least 30% below it. But gaps of the same size above the estimate are about as common (253 listings at least 43% above, against 276 at least 30% below), and the number of "cheap" listings depends strongly on the estimate used (276 listings at least 30% below the hedonic estimate, 476 below the benchmark). A price far below a model's estimate is therefore not by itself evidence of a bargain or of a hidden defect. Some extreme asking prices are simply errors, as the 1995 Audi A6 listed at 10,000 tenge shows, and an online marketplace has built automatic price-anomaly detection because inaccurate prices set by its sellers get published [13].

A small live check by the author illustrates information the model cannot see. On 10 October 2026 he reported on current kolesa.kz listings of the 2018 Toyota Camry: the average price was about 13 million tenge, and the prices ran from 7.5 to 17 million. In the 2025 dataset, the 12 listings of a 2018 Camry have a median of 12.2 million tenge (9.0 to 15.55 million), and the model's median estimate for them is 12.36 million. Of the cheapest car he found and another cheap one, the first, with a mileage of 260,000 km, had a check-engine light visible on a photo, and the second, with 185,000 km, a rear wheel that appeared to be a temporary spare. Mileage is a field of the listing, but the warning light and the spare wheel appear only in photos or in the full text of a listing, which this dataset does not contain, so the model cannot see them.

Limitations

The prices are asking prices, not the prices at which cars were sold. The dataset is a single snapshot of a few thousand listings whose collection date and method are not documented, so it may not represent all listings on the platform, and results for 2025 may not hold for other years. The file has no city, although region of sale was part of both earlier Kazakhstani studies [4, 8], no customs status, which Bekishev found to lower prices by 33% [8], and almost no drive information. The cleaning rules could only read the beginning of each listing, and they miss misspelled or Kazakh-language forms such as «не нахаду» and «жүрмей тұр». Mileage is missing in 485 listings and implausible in a few. Only one type of

model, a regularised linear regression, was tested; tree ensembles, which performed best in earlier studies [3, 4], might reduce the error further.

The original research question, approved by the supervisor on 1 October 2026, also asked what share of the listings priced far below the estimate is explained by information the seller discloses in the listing. This part was dropped, because the dataset contains only the first 121 characters (median) of each listing. Of the 276 listings at least 30% below the estimate, 145 have no seller text in that fragment (in 21 of them the fragment is complete, so the seller wrote nothing there), 101 have seller text that is cut off and 30 have complete seller text. A keyword search finds a reason for the low price (urgency, a defect or repair, or a legal problem such as transit plates or foreign registration) in 31 of the 276; the search skips negated words such as «без ДТП» or «не битая». Reading the 30 complete texts, which the author's research mentor did with an AI assistant, finds 4 more reasons that the keyword search misses («Аварииная», «Жүрмей тұр мотор жасату керек», «Строчна» and «Рус учет», a foreign registration), and no reason in the other 15. So 35 listings certainly give a reason in the text and 36 (the 15 complete texts and the 21 complete empty fragments) certainly give none; the share explained by the seller's text could therefore be anywhere between 12.7% and 87.0%, which is too wide to answer the question. The author's forecast of 7 October 2026, that about 70% of such listings would be explained by the reasons he had listed, lies inside this range and can be neither confirmed nor rejected with this file.

Conclusion

On 2,546 used-car listings from Kazakhstan, a hedonic regression on the fields a buyer can see in a listing estimated the asking price with a typical error of 16.7% and explained 86.2% of the variation in log price, better than the median for the same brand, model and year (22.5% and 60.3%); where listings of the same brand, model and year existed, their median was as precise as the model (12.9% vs 13.7%), and the model's advantage comes from cars without such comparables. Age and the brand-model pair carried most of the information, and each additional year of age was associated with a 6.8% lower asking price, almost the same as in the 2020 Kazakhstani estimate. The estimate is weakest for old cars (median error 29.9% for cars over 30 years old against 12.6% for cars up to 3 years old). About one listing in ten is priced at least 30% below the estimate, and gaps of the same size above the estimate are about as common, so a low price relative to the model is not by itself a sign of a bargain or of a hidden defect. Explaining such prices would need the full listing text, photos, the city and the customs status, which would make it possible to test the author's hypothesis directly.

Author contributions and use of AI tools. Kaiyrbek Zhanadil chose the topic, a price model for used-car listings in Kazakhstan, on 29 July 2026 from four directions proposed by his mentor; found Bekishev's study [8] on 17 September 2026, the studies [3, 4, 5] on 19 September, the studies [6, 7, 13] on 27 and 28 September, and the earlier paper in this journal [12] and the copy of [1] used here on 4 October, and summarised the studies he found for his mentor, mostly in his own words and in places with short summaries made by an AI tool; wrote the hypothesis on 24 September; drafted the rule for comparable cars, sent on 27 September; on 2 October chose and ranked the listing fields he expected to matter for the price (the model uses six of the eight he named, because drive is almost always empty in the file and customs status is absent, and three more); checked two other Kazakhstani car-listing websites as possible data sources on 7 October; wrote his forecast on 7 October; and reported a live check of 2018 Camry listings on 10 October (all dates in Astana time). The research question

was formulated by the author together with his research mentor; his research mentor narrowed it to its first half on 12 October 2026, after the dataset was chosen. His research mentor chose the dataset, set the rules for selecting and cleaning the listings, wrote and ran the analysis code, made the tables and figures, drew the conclusions and wrote this text in English, including the review of earlier work, which was written from the studies themselves, using an AI assistant (a large language model) for coding, reading the matched listing texts and drafting. No number was produced by the language model itself: every number in the article comes from the saved output of the analysis code, and every reference was opened and checked against its source.

Data availability. The analysis code is available from the author or his mentoring team on request; the data are available from [10].

Conflict of interest. The author declares no conflict of interest.

References:

1. Akerlof G.A. The Market for "Lemons": Quality Uncertainty and the Market Mechanism // The Quarterly Journal of Economics. 1970. Vol. 84, No. 3. P. 488-500. URL: https://www.jstor.org/stable/1879431

2. Lewis G. Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors // American Economic Review. 2011. Vol. 101, No. 4. P. 1535-1546. DOI: 10.1257/aer.101.4.1535

3. Fayyaz I., Ali G.G.M.N., Khairunnesa S.S. Advanced Feature Engineering and Machine Learning Techniques for High Accurate Price Prediction of Heterogeneous Pre-Own Cars // Vehicles. 2025. Vol. 7, No. 3. Art. 94. DOI: 10.3390/vehicles7030094

4. Barlybayev A., Sankibayev A., Kadyr Y., Amangeldy N., Sabyrov T. Predicting Used-Vehicle Resale Value in Developing Markets: Application of Machine Learning Models to the Kazakhstan Car Market // Ingénierie des Systèmes d'Information. 2023. Vol. 28, No. 5. P. 1237-1246. DOI: 10.18280/isi.280512

5. Sanjay R., Shiammala P.N. Used Car Price Prediction Using Machine Learning // International Journal of Creative and Open Research in Engineering and Management. 2026. Vol. 2, No. 5. P. 1-8. DOI: 10.55041/ijcope.v2i5.009

6. Vinnarasi F.S.F., Muthuvel A. Best Price Prediction for Pre-Owned Cars using Machine Learning Techniques // 2025 5th International Conference on Expert Clouds and Applications (ICOECA). Bengaluru: IEEE, 2025. P. 139-144. DOI: 10.1109/ICOECA66273.2025.00034

7. Zheng Y. Machine Learning Optimization and Challenges in Used Car Price Prediction // ITM Web of Conferences. 2025. Vol. 70. Art. 04032. DOI: 10.1051/itmconf/20257004032

8. Бекишев Р. Анализ ценообразования на автомобильном рынке Казахстана [Электронный ресурс] // AERC. 21.02.2020. URL: https://aerc.org.kz/kz/issledovanie/analiz-cenoobrazovaniya-na-avtomobilnom-rynke-kazahstana (accessed: 12.10.2026)

9. Rosen S. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition // Journal of Political Economy. 1974. Vol. 82, No. 1. P. 34-55. DOI: 10.1086/260169

10. mutashevdarkhan. Kolesa-Cars-2025 [Electronic resource]: dataset, version 1. 18.12.2025. URL: https://www.kaggle.com/datasets/mutashevdarkhan/kolesa-cars-2025 (accessed: 12.10.2026)

11. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning: with Applications in R. 2nd ed. New York: Springer, 2021. DOI: 10.1007/978-1-0716-1418-1

12. Mubarak B. Scalable Machine Learning Techniques for Fraud Detection in Financial Big Data // Central Asian Scientific Journal. 2026. No. 2(30), Vol. 2. P. 270-278. URL: https://cajournal.kz/vol/2026/2(30)_2.pdf (accessed: 12.10.2026)

13. Sarpal A., Kang Q., Huang F., Song Y., Wan L. A Marketplace Price Anomaly Detection System at Scale [Electronic resource] // arXiv. 2023. arXiv:2310.04367. DOI: 10.48550/arXiv.2310.04367

МАЗМҰНЫ – СОДЕРЖАНИЕ – CONTENT

ӘЛЕУМЕТТІК-ЭКОНОМИКАЛЫҚ ЖӘНЕ ҚҰҚЫҚ ҒЫЛЫМДАРЫ - СОЦИАЛЬНО-ЭКОНОМИЧЕСКИЕ И ЮРИДИЧЕСКИЕ НАУКИ - SOCIO-ECONOMIC AND LEGAL SCIENCES

Төлегенов Тимур Сейфуллаулы
ДИНАМИКА ТУРИСТСКИХ ПОТОКОВ И ЭФФЕКТИВНОСТЬ ИНФРАСТРУКТУРЫ ТУРКЕСТАНА: ОТ ЭКСТЕНСИВНОГО РОСТА К СМАРТ-УПРАВЛЕНИЮ 12

Alexander Gorevoy
WHICH INEQUALITY STATISTICS SURVIVE AN ADMINISTRATIVE BOUNDARY CHANGE? EVIDENCE FROM KAZAKHSTAN, 2010–2023 16

Rakhimov Djalil
DOES THE SHIFT TO CASHLESS PAYMENTS GO TOGETHER WITH FASTER GROWTH IN HOUSEHOLD CREDIT? EVIDENCE FROM THE REGIONS OF KAZAKHSTAN, 2023-2025 28

Kazhgerey Asan
HOW MANY FOOD BASKETS DOES AN AVERAGE WAGE BUY? WAGES AND FOOD PRICES IN THE 20 REGIONS OF KAZAKHSTAN, 2024-2026 39

Tagimova Alua
WHO WAS LEFT BEHIND? EDUCATION AND FINANCIAL INCLUSION IN KAZAKHSTAN AND THREE CENTRAL ASIAN NEIGHBOURS, 2021-2024 46

Wei Wei, Mynzhanova Gulzhakan, Niyetalina Gaukhar
DIGITAL–PHYSICAL INTEGRATION AND ECONOMIC PERFORMANCE IN INDUSTRIAL INTEGRATED ENERGY SYSTEMS: A SCENARIO-BASED ANALYSIS OF SMART GRID CONTROL, BATTERY ENERGY STORAGE, AND DIGITAL TWINS 53

Xu Fumin, Wei Wei, Ruziyeva Elvira, Moldabekova A.
FROM CULTURAL RESOURCES TO GLOBAL CONSUMER VALUE: THE BRIDGE–IMC MODEL AND THE INTERNATIONALIZATION OF BLACK MYTH: WUKONG UNDER THE BELT AND ROAD INITIATIVE 66

Ibraimkhan Alikhan
WHICH CRIMES CAN A 14-YEAR-OLD BE TRIED FOR? THE LIST IN ARTICLE 15 OF KAZAKHSTAN'S CRIMINAL CODE COMPARED WITH 30 OTHER JURISDICTIONS 82

Yerboluly Madi
DID IMPORTED CARS REALLY GET CHEAPER? CHINESE CARS AND THE AVERAGE CUSTOMS VALUE OF PASSENGER CARS IMPORTED INTO KAZAKHSTAN, 2015-2025 93

Салимбаев Ербол Аманжолович
АНАЛИЗ ИНВЕСТИЦИОННОЙ АКТИВНОСТИ КАК ИНДИКАТОР ИЗМЕНЕНИЙ ПОТЕНЦИАЛЬНОГО СПРОСА НА КОНСАЛТИНГОВЫЕ УСЛУГИ 101

Rambergenov Ramil
REGIONAL NEW-FIRM FORMATION IN UZBEKISTAN AFTER THE 2017 REFORMS: PERSISTENCE IN SCALE, CONVERGENCE IN INTENSITY, AND RECOMPOSITION TOWARD TRADE 109

Abdesh Dinara
FROM A WOMEN'S LEAD TO A MEN'S LEAD: THE GENDER GAP IN DIGITAL PAYMENTS IN KAZAKHSTAN AND THREE CENTRAL ASIAN NEIGHBOURS, 2021-2024 118

Amangeldi Aidana
DID SMALL BUSINESS IN KAZAKHSTAN REALLY SHRINK IN 2026? THE NEW TAX CODE, REGIONAL TAX RATES AND THE FALL IN THE NUMBER OF INDIVIDUAL ENTREPRENEURS 126

Cheng Che
BUILD, BUY, OR PARTNER? ENTRY MODE SELECTION AND RISK GOVERNANCE FOR WIND POWER INVESTMENT IN KAZAKHSTAN **135**

Wang Liang
FROM CAPACITY EXPANSION TO SYSTEM VALUE: COORDINATED DEVELOPMENT OF RENEWABLE ENERGY, GRID FLEXIBILITY AND ENERGY STORAGE IN KAZAKHSTAN **148**

Chen Jun
CONSTRUCTION-PERIOD CREDIT RISK MANAGEMENT FOR LARGE INFRASTRUCTURE PROJECT FINANCE IN KAZAKHSTAN **161**

Alniyaz Danial
HOW MUCH FUEL DOES A HYBRID REALLY SAVE? A SAME-MODEL COMPARISON OF HYBRID AND GASOLINE CARS IN US FUEL ECONOMY **175**

Assem Yergali, Maral Assanova
DIGITAL TRANSFORMATION AND SUSTAINABLE BUSINESS MODEL INNOVATION IN SMALL AND MEDIUM-SIZED ENTERPRISES: A SYSTEMATIC REVIEW OF DYNAMIC CAPABILITIES, VALUE CREATION, AND THE CREDIBILITY GAP IN EMERGING ECONOMIES **182**

Бейбіт Инара Алибекқызы, Едилов Саназар
ВОДА ИЗ ТУМАНА ИЛИ ВОДА ИЗ МОРЯ: ПОРОГОВАЯ СТОИМОСТЬ, ПРИ КОТОРОЙ ТУМАНОУЛОВИТЕЛИ КОНКУРЕНТОСПОСОБНЫ В МАНГИСТАУСКОЙ ОБЛАСТИ **195**

Zhenisbekova Aknur
DID A TENGE DEPOSIT OR THREE KAZAKHSTANI SHARES PROTECT SAVINGS FROM INFLATION? REAL RETURNS OF BANK DEPOSITS, KAZATOMPROM, KEGOC AND KASPI.KZ, 2020-2025 ..... **207**

Абилдаева Аселя
ОРГАНИЗАЦИОННАЯ ГОТОВНОСТЬ ВУЗОВ КАЗАХСТАНА К ВНЕДРЕНИЮ ЭНДАУМЕНТ-ФОНДОВ: КОНЦЕПТУАЛЬНАЯ РАМКА НА ОСНОВЕ ОБЗОРА ЛИТЕРАТУРЫ И АНАЛИЗА ДОКУМЕНТОВ **215**

Yang Juntao
DIGITAL ESG GOVERNANCE AND NEXT-GENERATION DIGITAL COMPETENCY AS DRIVERS OF SUCCESSION IN FAMILY AGRICULTURAL ENTERPRISES: A CONCEPTUAL FRAMEWORK ... **229**

Kairbek Mariya Islamkyzy, Kylbaev Erlan Serikovich
STATE SUPPORT FOR THE DIGITALIZATION OF MANUFACTURING IN THE REPUBLIC OF KAZAKHSTAN: AN ANALYSIS OF ITS COVERAGE OF MARKETING AND SALES ACTIVITIES **242**

Tilep Aiaru
WHEN DID IT EXPORTS TAKE OFF? ICT SERVICE EXPORTS OF KAZAKHSTAN, ARMENIA, GEORGIA AND THREE NEIGHBOURS, 2015-2024 ..... **255**

Anuarbek Murager
DID THE TENGE STOP FOLLOWING OIL? BRENT PRICES AND THE TENGE-DOLLAR RATE BEFORE AND AFTER FEBRUARY 2022 ..... **266**

Kaiyrbek Zhanadil
HOW ACCURATELY CAN OBSERVABLE LISTING DATA ESTIMATE FAIR PRICES FOR USED CARS IN KAZAKHSTAN? A HEDONIC MODEL OF 2,546 ONLINE LISTINGS **276**

Wen Yali
PERCEIVED ORGANIZATIONAL SUPPORT AND TURNOVER INTENTION AMONG KNOWLEDGE WORKERS IN A CHINESE PRIVATE TRAINING INSTITUTION: THE MEDIATING ROLE OF AFFECTIVE ORGANIZATIONAL COMMITMENT **289**

Құрманалиева Дана
ИСТОЧНИКИ ФИНАНСИРОВАНИЯ ИННОВАЦИОННЫХ ПРОЕКТОВ И КРИТЕРИИ ВЫБОРА ФИНАНСОВОЙ МОДЕЛИ **300**

ТЕХНИКАЛЫҚ ҒЫЛЫМДАР - ТЕХНИЧЕСКИЕ НАУКИ - TECHNICAL SCIENCE

Nurmakhan Almira
ADVANTAGES OF FIBER OPTIC COMMUNICATION OVER TRADITIONAL COPPER CABLES **307**

Kozhabai Bereke
SECURITY CHALLENGES IN IOT-BASED TELECOMMUNICATION NETWORKS **315**

Azamatuly Yerassyl
WHY AI-GENERATED CODE FAILS A BENCHMARK: A FAILURE ANALYSIS OF THREE OPEN LANGUAGE MODELS ON 30 HUMANEVAL+ TASKS **324**

Nurzhan Sanzhar
A FEW SHELLS, A FEW OPERATORS: WHERE LOW EARTH ORBIT IS CROWDED IN AN ALTITUDE-INCLINATION MAP OF 16,620 ACTIVE SATELLITES **333**

Askarbek Nurislam
CAN THE WEATHER TELL WHICH DAYS IN ASTANA WILL BE POLLUTED? CLASSIFYING PM2.5 EXCEEDANCE DAYS WITH MACHINE LEARNING AND CHECKING THE DATA AGAINST A GROUND MONITOR **339**

Iliyas Arman
SHOOTING ON THE MOVE: HOW MUCH AIM A ROBOT LOSES BY IGNORING ITS OWN VELOCITY, AND WHETHER A VIRTUAL-TARGET CORRECTION RECOVERS IT **350**

Ануар Әлімжан Бахытжанұлы, Сексенбаева Айша Курманбековна
СРАВНИТЕЛЬНЫЙ АНАЛИЗ СТАНДАРТОВ И МЕТОДИК ОБУЧЕНИЯ КИБЕРБЕЗОПАСНОСТИ: NICE FRAMEWORK, ECSF И ПРАКТИКО- ОРИЕНТИРОВАННЫЕ ТЕХНОЛОГИИ **356**

Mukhamedzhan Yerganat
GROWTH OF GITHUB DEVELOPER COMMUNITIES IN KAZAKHSTAN, UZBEKISTAN AND KYRGYZSTAN, 2020-2026: EVIDENCE FROM THE GITHUB INNOVATION GRAPH **366**

Kashkinov Alden
WHAT DECIDES HOW LONG A LANDER WORKS ON ANOTHER WORLD? SURFACE LIFETIMES OF 55 LANDERS ON THE MOON, VENUS, MARS, TITAN AND SMALL BODIES IN THE GCAT CATALOGUE, 1966-2025 ..... **378**

Kenesbek Miras
HOW MANY HERONIAN TRIANGLES HAVE A GIVEN LONGEST SIDE? A COMPLETE COUNT UP TO 2000 AND THE RIGHT TRIANGLES HIDDEN INSIDE **390**

Amirlan Seitkadyrov, Akezhan Omashev
AUTOMATIC MULTIPLE-CHOICE QUESTION GENERATION USING ARTIFICIAL INTELLIGENCE **401**

Idrissov Amirzhan
WHERE DO SMALL OPEN LANGUAGE MODELS BREAK ON OLYMPIAD MATHEMATICS? DOMAIN, DIFFICULTY AND MODEL SIZE ON OMNI-MATH **410**

Маматаева Рамина Адилгазыевна, Бидуәлі Алмат Алматұлы
ОБЗОР И СРАВНЕНИЕ ПОПУЛЯРНЫХ КРОССПЛАТФОРМЕННЫХ ФРЕЙМВОРКОВ ДЛЯ РАЗРАБОТКИ МОБИЛЬНЫХ ПРИЛОЖЕНИЙ **415**

Қасен Ғауһар Ғалымжанқызы, Алкеева Дана Мирасовна
КОМПЬЮТЕРЛІК КӨРУДІ ПАЙДАЛАНЫП,МЕХАНИКАЛЫҚ ДЕНЕ ЗАҢЫН АНЫҚТАУ АЛГОРИТМДЕРІН ӘЗІРЛЕУ **421**

Khalilzade Seymur Fazil oglu
ASSESSING THE IMPACT OF ELECTRIC VEHICLES ON ENERGY DEMAND IN HIGH-RISE RESIDENTIAL BUILDINGS **434**

Igilikov Adilkhan Igilikovich
PRE-TRAINED REPRESENTATIONS OF PRIOR-ADMISSION HISTORY IMPROVE ONE-YEAR, BUT NOT IN-HOSPITAL, MORTALITY PREDICTION IN ICU PATIENTS WITH CHRONIC DISEASE: A MIMIC-IV STUDY **449**

Nurgali Dias, Nurgali Miras, Akashov Aslan
PREDICTIVE TRAFFIC MANAGEMENT: EXPLORING SOLUTIONS TO TRAFFIC CONGESTION IN ASTANA, KAZAKHSTAN **461**

ПӘНДЕР АРАЛЫҚ ҒЫЛЫМДАР – МЕЖДИСЦИПЛИНАРНЫЕ НАУКИ – INTERDISCIPLINARY SCIENCES

Albiyev Asser
HOW MUCH OF AN ARTICLE COUNT IS HUMAN WORK? DECOMPOSING THE CREATION OF FIVE WIKIPEDIAS, 2010–2025 ..... **467**

Amangeldy Nurbolat
HOW MUCH OF KAZAKHSTAN'S URBAN BUILT STOCK PREDATES ITS SEISMIC CODES, AND WHERE DOES THAT STOCK SIT RELATIVE TO HAZARD? **477**

Torekhan Sabrina
WHERE ALMATY WALKS ITS OWN WAY: INFORMAL SHORTCUTS IN OPENSTREETMAP AND THE DETOURS THEY AVOID **483**

Ailan Salybekova
PARENTAL SUPPORT, PARENTAL CONTROL, STRESS AND EARLY ALCOHOL USE IN ADOLESCENT GIRLS: A NARRATIVE REVIEW WITH A FOCUS ON KAZAKHSTAN **488**

Yerboluly Madi
WHY IS KAZAKHSTAN'S GENDER GAP AT THE TOP OF CHESS SO SMALL? A BIRTH-COHORT COMPARISON ON THE FIDE RATING LIST **496**

Omashev Akezhan
REGIONAL HEALTH AND SOCIAL INDICATOR PATTERNS AND PCA- BASED MORTALITY MODELING IN KAZAKHSTAN, 2011–2024. A TWO- PART EXPLORATORY ECOLOGICAL ANALYSIS **501**

Мадиярова Жанел Биржановна
ДИПФЕЙК ТЕХНОЛОГИЯСЫНЫҢ ҚОҒАМДЫҚ ПІКІРГЕ ЫҚПАЛЫ ЖӘНЕ ОНЫ ҚҰҚЫҚТЫҚ РЕТТЕУ ТӘЖІРИБЕСІ **511**

Bolatov Sanzhar
HOW MUCH DOES KAZAKH SPEECH RECOGNITION LOSE WHEN THE MODEL GETS SMALLER? KAZAKH AGAINST 82 LANGUAGES IN THE PUBLISHED WHISPER RESULTS **521**

Ilya Tsoy
WHO PAYS MORE TOKENS FOR THE SAME SENTENCE? RUSSIAN, KOREAN AND ENGLISH IN FOUR LANGUAGE MODEL TOKENIZERS ON 1,012 FLORES-200 SENTENCES **530**

Assan Ramazan
WHICH NAME DO READERS STILL USE? FORMER NAMES OF RENAMED KAZAKHSTANI CITIES IN RUSSIAN AND KAZAKH WIKIPEDIA PAGE VIEWS, 2015–2026 ..... **543**

Nigmadilova Zhasmin
WHAT REPLACES PARABENS? THE PARABEN-FREE LABEL ACROSS PRODUCT CATEGORIES IN THE OPEN BEAUTY FACTS DATABASE **556**

ӘЛЕУМЕТТІК ЖӘНЕ ГУМАНИТАРЛЫҚ ҒЫЛЫМДАР - СОЦИАЛЬНО-ГУМАНИТАРНЫЕ НАУКИ - SOCIAL SCIENCES AND HUMANITIES

Мұқатаева Айзат Амантайқызы, Стамбекова Жанат Байғазықызы, Слямханова Гүлім Болатқызы, Қызырбек Аяулым Ерболқызы
ЖАСАНДЫ ИНТЕЛЛЕКТ ДӘУІРІНДЕГІ ҒЫЛЫМИ ТАНЫМ: ПОППЕР, КУН ЖӘНЕ ЛАКАТОС ИДЕЯЛАРЫ ТҰРҒЫСЫНАН ТАЛДАУ **561**

Мұқатаева Айзат Амантайқызы, Сапарханова Мадина Төлегенқызы, Окенова Аружан Муратхановна, Тұрсынғазыева Ақерке Ерболатқызы
ШӘКӘРІМ ФИЛОСОФИЯСЫНДАҒЫ ЕРКІНДІК ЖӘНЕ МОРАЛЬДЫҚ ТАҢДАУ МӘСЕЛЕСІ **571**

Komekova Saida
DO UNHAPPY REVIEWERS MIX IN MORE RUSSIAN? STAR RATINGS AND KAZAKH-RUSSIAN CODE-SWITCHING IN 175,002 ONLINE REVIEWS **580**

Zhumagali Yerbol
DOES A BIGGER MODEL CLOSE THE GAP? KAZAKH-LANGUAGE ACCURACY ON KAZAKHSTAN-SPECIFIC VERSUS UNIVERSAL SCHOOL SUBJECTS IN SMALL OPEN-WEIGHT LANGUAGE MODELS **593**

Abdrakhmanova Madina
DOES A MEDAL BUY ATTENTION? WIKIPEDIA VIEWS OF KAZAKHSTAN'S ATHLETES BEFORE AND AFTER THE PARIS 2024 OLYMPIC **597**

ПЕДАГОГИКАЛЫҚ ЖӘНЕ ПСИХОЛОГИЯЛЫҚ ҒЫЛЫМДАР – ПЕДАГОГИЧЕСКИЕ И ПСИХОЛОГИЧЕСКИЕ НАУКИ – PEDAGOGICAL AND PSYCHOLOGICAL SCIENCES

Kaldarbek Mergen
FAMILY OR SCHOOL? SOCIOECONOMIC COMPOSITION AND MATHEMATICS ACHIEVEMENT IN KAZAKHSTAN, PISA 2022 AND 2025 ..... **605**

Zhenis Adilzhan
WHICH OLYMPIAD PROBLEMS DO STUDENTS TAKE TO THE BOARD? CHOICES AND DEFENCES IN THE 8TH-GRADE LEAGUE OF KARAGANDY MATHBATTLES 2026 ..... **612**

Базарбаева Айнур Бектуровна
ДЕТСКИЙ ФОЛЬКЛОР И ПОДВИЖНЫЕ ИГРЫ КАК ИНСТРУМЕНТЫ ФОРМИРОВАНИЯ ОСНОВ ПРАВОВОЙ КУЛЬТУРЫ У ДЕТЕЙ СТАРШЕГО ДОШКОЛЬНОГО ВОЗРАСТА **619**

БИОЛОГИЯ ҒЫЛЫМДАР - БИОЛОГИЧЕСКИЕ НАУКИ - BIOLOGICAL SCIENCES

Bazhenova Diana
GROWTH DYNAMICS OF STERLET (Acipenser ruthenus L.) REARED IN A RECIRCULATING AQUACULTURE SYSTEM **622**

ТУРИСТИЧЕСКИЕ И СПОРТИВНЫЕ НАУКИ - ТУРИСТІК ЖӘНЕ СПОРТ ҒЫЛЫМДАРЫ - TOURISM AND SPORTS SCIENCES

Төлеп Жазира Маратқызы
ГАСТРОНОМИЯЛЫҚ ФЕСТИВАЛЬДЕРДІҢ ТУРИСТІК ДЕСТИНАЦИЯ ТАРТЫМДЫЛЫҒЫН ҚАЛЫПТАСТЫРУДАҒЫ РӨЛІ МЕН ДАМУ ҮРДІСТЕРІ **629**

ЕСТЕСТВЕННЫЕ НАУКИ - ЖАРАТЫЛЫСТАНУ ҒЫЛЫМДАРЫ - NATURAL SCIENCES

Нургельдина Асель Ермековна, Серикбаева Жанерке Алимжановна
ШЕКАРАСЫЗ ФИЗИКА: ӘР ЕЛДЕГІ ФИЗИКАЛЫҚ БІЛІМДІ САЛЫСТЫРУ **636**

Torekhan Sabrina
IS ALMATY A 15-MINUTE CITY? WALKING ACCESS TO SIX EVERYDAY PLACES FROM 170,838 BUILDINGS IN THE CITY'S EIGHT DISTRICTS **643**

Онгарбек Махабат Саятқызы
CU-PT ЖӘНЕ NI-PT ҚОРЫТПАЛАРЫНЫҢ ФИЗИКАЛЫҚ ҚАСИЕТТЕРІНЕ РЕТТЕЛУ ҮДЕРІСІНІҢ ЖӘНЕ ҚҰРЫЛЫМДЫҚ ӨЗГЕРІСІНІҢ ӘСЕРІН РЕТТЕУ **651**

Нургельдина Асель Ермековна, Әбдуәлі Мерей Ералықызы
ҮЙКЕЛІС КҮШІНІҢ ТАБИҒАТТАҒЫ РӨЛІ **659**

Ossipov Vladislav
MANY REPORTS, WEAK EVIDENCE? EVIDENCE LEVELS FOR MICROBIAL PLASTIC DEGRADATION IN THE 2,535 RECORDS OF PLASTICDB **664**

Нургельдина Асель Ермековна, Кабдулаева Жанель Батырбековна
РАДИОАКТИВТІЛІК. АЛЬФА, БЕТА, ГАММА СӘУЛЕЛЕРІНІҢ ТЕГІ МЕН ЗАҢДЫЛЫҚТАРЫ. ЯДРОЛЫҚ РЕАКЦИЯ ЖӘНЕ ЖАРТЫЛАЙ ЫДЫРАУ ПЕРИОДЫ **673**

Kashkinov Alden
DOES AN ACTIVE SUN BRING SATELLITES DOWN SOONER? SOLAR ACTIVITY AND THE RE-ENTRY OF UNCONTROLLED OBJECTS IN LOW EARTH ORBIT, 1990-2025 ..... **678**

Demin Mark
HOW MUCH LESS OXYGEN CAN THE WARMING CASPIAN SEA HOLD? THE LOSS OF OXYGEN SOLUBILITY IN THE NORTHERN, MIDDLE AND SOUTHERN CASPIAN, 1982-2022 ..... **684**