The American Journal of Applied Sciences

ISSN 2689-0992 | Open Access

TYPE Original Research
PAGE NO. 77-84
DOI 10.37547/tajas/Volume07Issue11-09

Check for updates

OPEN ACCESS

SUBMITTED 29 October 2025
ACCEPTED 12 November 2025
PUBLISHED 28 November 2025
VOLUME Vol.07 Issue 11 2025

CITATION
Oleksii Fonin. (2025). Using AI Forecast for Identification Injuries and Fatalities in Equine Industry. The American Journal of Applied Sciences, 7(11), 77–84. https://doi.org/10.37547/tajas/Volume07Issue11-09

COPYRIGHT
© 2025 Original content from this work may be used under the terms of the creative common's attributes 4.0 License.

Using AI Forecast for Identification Injuries and Fatalities in Equine Industry

Oleksii Fonin
Founder and CEO of 49 IT Group, Senior Engineering Manager for HISA Portland, Oregon, USA

Abstract: The article presents a broad analysis of methods for applying artificial intelligence to predict injuries and fatalities in the Thoroughbred industry. The study is based on an interdisciplinary approach that incorporates elements of veterinary medicine, sensor technologies, radiomics, image analysis, and the processing of clinical records. Particular attention is given to a comparative-analytical review of modern models, including computer vision techniques, radiomic analysis (μCT), machine learning algorithms, and deep neural network architectures (CNN, ANN), as well as the use of inertial measurement units (IMU) for quantitative risk stratification. The limitations of individual approaches are identified: small sample sizes and low classification stability for eye regions in computer vision systems, high cost and limited accessibility of radiomics, and the dependence of sensor technologies on track surface and conditions. The integration of domain transfer methods, facial expression analysis, and NLP services for unstructured medical data processing is emphasized. The impact of multimodal integration of images, sensor metrics, and clinical records on improving the accuracy and robustness of predictions is examined. The need to shift from universal monitoring strategies to targeted control of high-risk groups is substantiated, enabling proactive equine health management and reducing fatality rates in the racing industry. The article proposes an original classification of promising development directions, including complex architectures, the use of Retrieval Augmented Generation to link sensor and clinical data, and the implementation of integrated platforms in the practice of insurance, breeding, and injury prevention. The material will be of interest to researchers in veterinary medicine, sports analytics, artificial intelligence, and

digital technologies, as well as to specialists engaged in the development and implementation of risk prediction systems in the equine industry.

Keywords: Artificial intelligence, equine industry, injury prediction, fatal musculoskeletal injuries, computer vision, inertial measurement units, multimodal data integration.

Introduction

The modern performance and racehorse industry is undergoing radical changes in the context of digital transformation, related to the implementation of artificial intelligence, the development of sensor technologies, and the growth in the volume of data on the physiological state of animals. The increasing complexity of risk assessment models and the shift from subjective visual observation to algorithmically controlled monitoring systems are creating new demands for the accuracy and reliability of injury predictions. The problem of early detection of precursors to fatal injuries and the objective diagnosis of pain is becoming particularly relevant, as it is directly related to the safety of horses, the sustainability of the sports industry, and public trust in equestrian sports.

Against the backdrop of the rapid development of data analysis technologies, there is a growing interest in methods that allow for the objective recording of signs of pain, laminitis, and hidden pathologies of the musculoskeletal system, as well as for predicting the probability of fatal injuries [3]. The scientific and applied literature increasingly raises the issue of integrating computer vision, sensor systems, and machine learning algorithms. Both classical classification models and more complex deep learning architectures based on the processing of images, biomechanical data, and clinical records are being considered. These studies are becoming particularly significant in the context of growing demands for animal welfare and the need to reduce the socio-economic risks associated with high mortality rates at racetracks.

An example of a comprehensive approach is solutions that combine computer vision, radiomic methods, inertial measurement sensors, and medical data analysis services [6]. Such technologies make it possible to identify deviations in movement, record micro-changes in facial expressions, compare sensory indicators with clinical protocols, and form individual risk forecasts for each horse. Such systems are already demonstrating a high level of accuracy in practical conditions, but they require adaptation to the specifics of veterinary medicine and the development of standards for interpreting the results.

The objective of this study is to conduct a systematic analysis of the methods for applying artificial intelligence to predict injuries and fatal outcomes in the racehorse industry, to classify existing approaches, to identify their strengths and weaknesses, and to propose directions for integrating multi-modal data (images, sensors, clinical records) into a unified risk prediction model.

Materials and Methods

The methodological foundation of this study is formed at the intersection of veterinary medicine, artificial intelligence technologies, and applied predictive analytics. The interdisciplinary nature of the topic necessitated the use of a theoretical analysis of modern scientific literature dedicated to the issues of automated pain detection, injury prediction, and the application of sensor systems in the performance and racehorse industry.

The study utilized sources covering experimental works on pain recognition based on computer vision and applied developments of algorithms for predicting fatal injuries based on sensor technologies. The work of Anderson K. [1] validated inertial measurement units (IMUs) for monitoring the behavior of horses in stall conditions, which formed the basis for the further application of sensory data in risk prediction. The study by Asti V. [2] made it possible to assess the accuracy of gait registration and the identification of biomechanical characteristics in different horse breeds, confirming the applicability of IMUs for early diagnosis tasks. The work of Basran P. [3] contributed to the methodology of analysis using radiomics and micro-computed tomography for predicting sesamoid bone fractures in racehorses.

A significant direction has been the development of algorithms based on machine learning. The study by Bogossian P. [4] showed that the use of models for analyzing stride and movement characteristics allows for the prediction of the probability of enforced rest and the career termination of horses. The application of domain transfer methods for recognizing orthopedic pain in video data was presented in the work of Broomé S. [5], where low-intensity forms of pain, difficult for an observer to distinguish, were recorded. The review by Crecan C. [6] systematized the experience of applying sensor technologies in gait analysis and emphasized their importance for the objective control of risks.

A separate block of research is devoted to the analysis of facial reactions. The work of Feighelstein M. [7] proposed algorithms for the automatic recognition of the emotional states of horses, confirming the possibility of detecting pain and stress through facial signs. The study by Kim S. M. [8] developed a deep learning model for classifying facial expressions into four states (rest, pain, exercise, shoeing), which expanded the possibilities for an objective assessment of animal welfare. The work of Lencioni G. C. [9] became the basis for building convolutional neural networks capable of classifying pain levels according to the Horse Grimace Scale (HGS), showing the potential of automated video systems. The final element of the methodological block is the study by Mc Sweeney D. [10], where, based on an analysis of 28,481 starts of 11,834 racehorses equipped with IMU sensors, it was shown that risk scores on a scale of 1 to 6 correlate exponentially with the probability of fatal injuries.

Thus, the methodological strategy of the research is based on a comprehensive analysis of sources that reflect the full spectrum of modern approaches, from computer vision and radiomics to the application of IMU sensors and machine learning algorithms. This allowed for the systematization of existing methods and the identification of directions for the further development of prognostic models in the racehorse industry.

Results

The application of deep neural network models in veterinary diagnostics has made it possible to form objective approaches to pain recognition in horses. The work of Lencioni G. C. [9] analyzed the diagnostic value of individual facial regions. It was established that different zones have different prognostic significance. The most reliable indicator was the ears, while the eyes proved to be the least informative. Additionally, the effectiveness of integrating features into artificial neural networks was tested, which made it possible to obtain stable results for both multi-level and binary classification. Table 1 examines the distribution of accuracy across key categories and facial regions, which allows for a comparison of the effectiveness of individual models and confirms the heterogeneity of the diagnostic value of the features.

Table 1. Comparative accuracy of AI models for equine facial expression analysis (Compiled by the author based on sources: [8, 9])

Categories / Region	Sample size	Accuracy (%)
Eyes–nose–ears detection (training)	–	98.75
Eyes–nose–ears detection (validation)	–	81.44
Eyes–nose–ears detection (testing)	–	88.10
Eyes–nose–ears detection (average)	–	89.43
Ears	2,379 images	90.3
Eyes	1,436 images	65.5
Mouth and nostrils	1,035 images	74.5
ANN classifier (3 levels)	120 images	75.8
ANN classifier (2 levels)	120 images	88.3

A comparison of the results shows that individual facial zones have different degrees of informativeness, and the integration of features allows for an increase in the stability of the final classification. The most promising are complex models that combine local regions with neural network algorithms, which creates a basis for the further implementation of automated pain monitoring systems in practice.

Modern research in the field of preventing fatal injuries in horses is increasingly turning to data obtained using inertial sensors. The work of Anderson K. [1] demonstrated the validity of IMU technology for recording behavioral patterns in stall conditions, which became the basis for transferring the methodology to a sports context. The study by Asti V. [2] conducted a comprehensive gait assessment using IMUs in Thoroughbred horses, where the high sensitivity of the sensors to micro-deformations of movement was confirmed. These results are directly related to the conclusions of Crecan C. [6], who emphasizes that sustainable prediction is possible only with a combination of quantitative biomechanics and machine learning algorithms.

The research by Bogossian P. [4], which identified predictors of enforced rest and premature culling of horses based on an analysis of stride characteristics, made a special contribution to the development of applied prediction models. These approaches allow sensory data to be considered as a diagnostic tool and a source of prognostic information about the long-term athletic career of animals. A similar perspective is reflected in the work of Basran P. [3], where radiomic models based on µCT were used to assess the risk of catastrophic fractures. Despite the difference in methods, both directions demonstrate a single trend—the need for early detection of pathologies to prevent a fatal outcome. Table 2 examines the distribution of the probability of FMI depending on the risk level. The data from McSweeney D. [10] show an exponential nature of the increase in odds when moving from minimum to maximum values of the integral risk score.

Table 2. Probability of fatal musculoskeletal injuries in Thoroughbreds depending on risk score (Compiled by the author based on source: [10])

Risk score	% of starts	FMI %	OR (95% CI)
1	64.5	0.12	1.0
2	17.1	0.08	1.21
3	10.4	0.37	3.99
4	5.3	0.46	5.38
5	2.4	1.2	8.76
6	0.4	4.2	44.6

Based on the results, the analysis of Table 2 reveals a clear non-uniformity in the distribution of fatal cases. The vast majority of starts fall into the minimum risk levels (risk score 1–2), where the frequency of FMI remains extremely low. However, it is the small cohort with the maximum risk score of 6 that accumulates the highest probability of fatal injuries, which confirms the exponential nature of the relationship between the risk level and the outcome.

The structure of Table 2 demonstrates several key parameters. The "% of starts" column reflects the proportion of races corresponding to each risk category, allowing for an assessment of the prevalence of the respective profiles in the sample. The "FMI %" indicator shows the proportion of fatal injuries within a given risk group, demonstrating an increase in the frequency of cases as the risk score increases. Finally, the "OR (95% CI)" represents the odds ratio, showing how many times higher the probability of FMI is compared to the baseline group (risk score 1). Particularly noteworthy is the jump in values in the interval between levels 5 and 6. This reflects a sharp increase in the probability of a fatal outcome upon reaching extreme risk values.

This distribution underscores that universal monitoring strategies are less effective compared to the targeted control of a small part of the population classified into the most vulnerable categories. The inclusion of IMU sensor data in prognostic models allows for the early identification of these groups, which opens up the possibility of early intervention and prevention. At the same time, a comparative analysis with the results of other studies confirms that the effectiveness of prediction increases when combining sensor technologies, facial expression analysis, and radiomics, forming the basis for comprehensive systems for preventing fatal injuries in the equestrian sports industry.

Discussion

A systematic analysis of the various applications of artificial intelligence in the equestrian sports industry shows that each has both unique advantages and significant limitations. First and foremost, computer vision and deep learning methods have proven their effectiveness in tasks of pain detection in horses. The study by Lencioni G. C. [9] showed that classification based on ears achieves high accuracy, while the eyes remain the least reliable indicator. Similar results were demonstrated in the work of Kim S. M. [8], where the use of ResNet and EfficientDet made it possible to achieve a stable level of accuracy on a sample of 749 horses. Additional confirmations of the applicability of the approach are contained in the studies by Broomé S. [5], Feighelstein M. [7], and Asti V. [2], which allows CV/Deep Learning methods to be considered as a basis for the objectification of visual pain diagnosis. At the same time, a weakness of these approaches remains the limited size of the samples and a reduction in the stability of classification for individual facial regions.

Radiomics has shown potential for early diagnosis. The study by Basran P. [3] using µCT made it possible to identify hidden structural anomalies associated with catastrophic fractures. However, the high cost and limited access to the equipment make the method difficult to scale in everyday practice. The application of methods for transferring models between subject areas has seen interesting development. The study by Broomé S. [5] showed that transferring models between different types of video can expand their field of application. However, classification errors when working with low-quality recordings limit the reliability of the results.

Sensor systems have demonstrated a fundamentally different level of scalability. The study by McSweeney D. [10] revealed an exponential relationship between the risk score and the probability of a fatal injury. As shown in Table 2, the probability of FMI at the maximum risk score value increased compared to the minimum risk group. The data confirm that sensor technologies allow for the quantitative stratification of the population and the establishment of monitoring priorities. At the same time, the scope of the study was limited to specific tracks and depended on the characteristics of the surface, which reduces the universality of the method.

In recent years, a promising direction has been the analysis of unstructured medical data using NLP services. The use of Amazon Comprehend Medical and Bedrock allows for the identification of key medical terms from the texts of veterinary records, and the use of RAG mechanisms increases relevance through the integration of a specialized knowledge base. A limitation, however, remains the absence of unified coding systems for veterinary practice and a bias towards human terminology. Table 3 summarizes the main advantages and limitations of the analyzed approaches.

Table 3. Comparative advantages and limitations of AI approaches (Compiled by the author based on sources: [1, 3, 4, 10])

Approach	Advantages	Limitations
CV/Deep Learning	High accuracy, pain objectification	Small datasets, low accuracy for eyes
Radiomics (µCT)	Early fracture diagnostics	Limited access, high cost
Domain Transfer (video)	Model transfer across contexts	Errors with low-quality video
IMU-sensors	Large datasets, OR = 44.6 (risk score 6)	Limited to certain tracks, surface dependence
NLP (Comprehend, Bedrock)	Structuring records, RAG-search	No unified veterinary coding, human bias

A comparative analysis confirms the need to integrate different directions. Computer vision methods provide a tool for daily pain diagnosis, sensor systems provide risk stratification at the population level, and radiomics and NLP technologies allow for an expansion of the depth of analysis through medical data and images. In aggregate, these approaches form the basis for comprehensive prediction models that can significantly reduce the level of fatal injuries and improve the quality of veterinary support in equestrian sports.

Modern research convincingly shows that the isolated application of individual artificial intelligence technologies has limited potential. Computer vision methods have demonstrated high accuracy in identifying pain expressions, especially in the area of the ears and muzzle [4], but the problem of small sample sizes and low stability when analyzing the eyes remains. Sensor systems have, for the first time, made it possible to build quantitative risk models using large-scale samples, but their coverage depends on the type of surface and the availability of tracks. Radiomics has confirmed its diagnostic value in assessing hidden pathologies [2], but high costs and limited access hinder its implementation in practice. These observations point to the need for a multi-modal combination of technologies.

The integration of CV, IMU, and NLP can overcome the limitations of each direction. Facial expression analysis systems provide the opportunity for continuous pain monitoring in stall and training base conditions, while sensors record load and micro-deformations of movement in real time [6]. Their combination allows for the formation of a comprehensive risk profile, where subjective signs of pain are supplemented by objective biomechanical parameters. Additionally, the inclusion of NLP services opens up the prospect of processing unstructured clinical records. The use of Amazon Comprehend Medical provides for the automatic extraction of medical terms and diagnoses from the texts of veterinary charts, and the application of Bedrock and Retrieval Augmented Generation allows for linking sensory data with clinical protocols and knowledge bases, forming enriched diagnostic models.

Of particular interest is the possibility of the practical application of integrated systems in various segments of the industry. In horse insurance, such technologies can be used for dynamic risk assessment: data on biomechanical load and clinical records allow for the formation of individual insurance products based on the actual condition of the animal. In breeding work, the use of complex models creates the prerequisites for selecting lines with a lower predisposition to injuries, which increases the resilience of populations. In injury prevention, multi-component systems can serve as an early warning tool: the combination of computer vision methods, sensory control, and analysis of textual data allows for the identification of weak signals that individually do not reach statistical significance, but in aggregate indicate a high risk.

Thus, the prospects for development lie in the creation of an integrated platform where various AI technologies will function as complementary modules. This approach is capable of increasing the accuracy of predictions and changing the very paradigm of veterinary support, shifting it from reactive treatment to proactive prevention. This is confirmed by data from studies in the fields of sensor technology, movement analysis, computer vision, and radiomics, which in aggregate form the basis for the transition to new generation systems.

Conclusion

The study conducted has made it possible to systematically analyze the methods of applying artificial intelligence for predicting injuries and fatal outcomes in the racehorse industry, to identify their strengths and weaknesses, and to substantiate the need for a transition to complex models capable of combining various data sources. It has been established that traditional approaches based on subjective observation and limited physiological indicators are losing their effectiveness in the face of the growing complexity of movement biomechanics, the increasing volume of clinical data, and the rising demands for prediction accuracy.

A comparative analysis showed that the greatest prognostic value is demonstrated by architectures that combine computer vision methods with neural network algorithms, sensor technologies, and statistical models of movement analysis. Such solutions make it possible to record both micro-signals of pain, expressed in the facial expressions of horses, and objective biomechanical parameters that reflect hidden pathologies of the musculoskeletal system. The application of radiomics confirmed the importance of visualization data for the early diagnosis of fractures, and the use of IMU sensors, for the first time, provided the opportunity for quantitative risk stratification at the population level. Particularly important was the confirmation of the exponential relationship between the value of the integral risk score and the probability of a fatal outcome, which demonstrates the potential of sensor technologies as a tool for practical prevention.

Separate attention was paid to the institutional and technological factors that affect the reliability of the forecasts. It was shown that the absence of unified coding standards in veterinary medicine limits the possibilities of integrating textual data, and the bias of existing NLP services towards human terminology reduces the accuracy of the analysis of veterinary records. At the same time, the development of services for processing unstructured data opens up the prospect of combining sensory indicators with clinical protocols and literary knowledge bases. This direction forms the basis for building Retrieval Augmented Generation systems capable of linking biomechanics with clinical information and increasing the accuracy of predictions.

The most important result of comparing the different approaches is the conclusion that universal monitoring strategies are less effective than the targeted control of groups with the maximum level of risk. The application of multi-component models allows for a transition from recording consequences to predicting the probability of events, which ensures proactive risk management and a reduction in mortality. This approach has applied significance for the sports industry and for insurance, breeding, and preventive veterinary medicine, where prediction becomes part of the strategy for managing populations.

Thus, the modern system for predicting injuries and fatal outcomes in the racehorse industry requires a transition from isolated methods to integrated platforms that combine visual, sensory, and textual data. The prospects for further research are related to the creation of unified coding standards for veterinary medicine, the expansion of the volume of samples, the application of hybrid deep learning architectures, and the formation of a regulatory framework for the implementation of algorithmic systems in the industry. It is the multidisciplinary integration and adaptability of models that are becoming the key factors in increasing the accuracy of predictions and the resilience of the equestrian industry.

References

1. Anderson, K., Morrice-West, A. V., Walmsley, E. A., Fisher, A. D., Whitton, R. C., & Hitchens, P. L. (2023). Validation of inertial measurement units to detect and predict horse behaviour while stabled. Equine Veterinary Journal, 55(6), P. 1128–1138. https://doi.org/10.1111/evj.13909
2. Asti, V., Ablondi, M., Molle, A., Zanotti, A., Vasini, M., & Sabbioni, A. (2024). Inertial measurement unit technology for gait detection: A comprehensive evaluation of gait traits in two Italian horse breeds. Frontiers in Veterinary Science, 11, Article 1459553. https://doi.org/10.3389/fvets.2024.1459553
3. Basran, P. S., McDonough, S., Palmer, S., & Reesink, H. L. (2022). Radiomics modeling of catastrophic proximal sesamoid bone fractures in Thoroughbred racehorses using µCT. Animals, 12(21), 3033. https://doi.org/10.3390/ani12213033
4. Bogossian, P. M., Nattala, U., Wong, A. S. M., & others. (2024). A machine learning approach to identify stride characteristics predictive of musculoskeletal injury, enforced rest and retirement in Thoroughbred racehorses. Scientific Reports, 14, Article 28967. https://doi.org/10.1038/s41598-024-79071-1
5. Broomé, S., Ask, K., Rashid-Engström, M., Haubro Andersen, P., & Kjellström, H. (2022). Sharing pain: Using pain domain transfer for video recognition of low grade orthopedic pain in horses. PLOS ONE, 17(3), e0263854. https://doi.org/10.1371/journal.pone.0263854
6. Crecan, C. M., & Peștean, C. P. (2023). Inertial sensor technologies—their role in equine gait analysis: A review. Sensors, 23(14), 6301. https://doi.org/10.3390/s23146301
7. Feighelstein, M., Riccie-Bonot, C., Hasan, H., Weinberg, H., Rettig, T., Segal, M., Distelfeld, T., Shimshoni, I., Mills, D. S., & Zamansky, A. (2024). Automated recognition of emotional states of horses from facial expressions. PLOS ONE, 19(7), e0302893. https://doi.org/10.1371/journal.pone.0302893
8. Kim, S. M., & Cho, G. J. (2023). Analysis of various facial expressions of horses as a welfare indicator using deep learning. Veterinary Sciences, 10(4), 283. https://doi.org/10.3390/vetsci10040283
9. Lencioni, G. C., de Sousa, R. V., de Souza Sardinha, E. J., Corrêa, R. R., & Zanella, A. J. (2021). Pain assessment in horses using automatic facial expression recognition through deep learning-based modeling. PLOS ONE, 16(10), e0258672. https://doi.org/10.1371/journal.pone.0258672

10. Mc Sweeney, D., Wang, Y., Palmer, S. E., Holmströem, M., Donohue, K. D., Farnsworth, K. D., Sanz, M. G., Lambert, D. H., & Bayly, W. M. (2025). Thoroughbreds deemed to be most at risk by inertial measurement unit sensors suffered a fatal musculoskeletal injury at a higher rate than other racehorses. Journal of the American Veterinary Medical Association. Advance online publication. https://doi.org/10.2460/javma.25.04.0268