

ΕΛΛΗΝΙΚΟ ΜΕΣΟΓΕΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΟΝΩΝ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ

ΤΜΗΜΑ ΔΙΟΙΚΗΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

<ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΣΤΟ ΧΩΡΟ ΤΟΥ ΑΘΛΗΤΙΣΜΟΥ>

Ονοματεπώνυμο Σπουδαστή: ΒΑΣΙΛΙΚΗ-ΕΛΕΝΗ ΣΙΛΒΑ-ΓΙΑΝΝΑΚΑΚΗ

Αρ. Μητρώου Σπουδαστή: ΔΕ20269

Επιβλέπων Καθηγητής: ΙΩΑΝΝΗΣ ΚΑΠΑΝΤΑΪΔΑΚΗΣ

2024-2025

HELLENIC MEDITERRANEAN UNIVERSITY

SCHOOL OF MANAGEMENT AND ECONOMICS SCIENCE

DEPARTMENT OF MANAGMENENT SCIENCE AND TECHNOLOGY

DIPLOMA THESIS

<DATA ANALYSIS IN THE FIELD OF SPORTS>

Student : Vasiliki-Eleni Silva-Giannakaki 20269

Supervisor : Ioannis Kapantaidakis

2024-2025

Υπεύθυνη Δήλωση : Βεβαιώνω ότι είμαι συγγραφέας αυτής της πτυχιακής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην πτυχιακή εργασία. Επίσης έχω αναφέρει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επίσης βεβαιώνω ότι αυτή η πτυχιακή εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για τις απαιτήσεις του προγράμματος σπουδών του Τμήματος Διοικητικής Επιστήμης και Τεχνολογίας του ΕΛ.ΜΕ.ΠΑ.

ΠΙΝΑΚΑΣ ΠΕΡΙΟΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ 5

ABSTRACT 6

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ 7

ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ 13

ΚΕΦΑΛΑΙΟ 3: ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΕΡΓΑΛΕΙΑ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ ΣΤΟΝ ΑΘΛΗΤΙΣΜΟ . 29

ΚΕΦΑΛΑΙΟ 4: CASE STUDIES 43

ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ 50

ΒΙΒΛΙΟΓΡΑΦΙΑ 52

ΠΕΡΙΛΗΨΗ

Η παρούσα εργασία εξετάζει τη σημασία και τις εφαρμογές της ανάλυσης δεδομένων στον χώρο του αθλητισμού, αναδεικνύοντας τον ρόλο της τεχνολογίας και της επιστήμης στη βελτίωση της απόδοσης, της στρατηγικής και της υγειονομικής διαχείρισης των αθλητών. Αρχικά παρουσιάζονται οι βασικές έννοιες της ανάλυσης δεδομένων, οι τύποι των δεδομένων που συλλέγονται (στατιστικά παικτών, φυσικής κατάστασης, βιομετρικά, ψυχολογικά) και οι κύριες πηγές άντλησής τους (wearables, αισθητήρες, βίντεο, εργομετρικά τεστ). Στη συνέχεια, αναλύονται οι μέθοδοι συλλογής, επεξεργασίας και στατιστικής αξιοποίησης των δεδομένων με τη χρήση σύγχρονων εργαλείων όπως το Excel, η Python, η R και η SQL. Ιδιαίτερη έμφαση δίνεται στην εφαρμογή αλγορίθμων μηχανικής μάθησης για την πρόβλεψη επιδόσεων, την πρόληψη τραυματισμών και την αγωνιστική στρατηγική. Μέσω μελετών περίπτωσης, αποδεικνύεται πως η ανάλυση δεδομένων δεν αποτελεί απλώς υποστηρικτικό εργαλείο, αλλά αναπόσπαστο στοιχείο της αθλητικής επιστήμης και διοίκησης. Τελικά, η εργασία καταλήγει στο συμπέρασμα ότι ο αθλητισμός του 21ου αιώνα βασίζεται σε μεγάλο βαθμό στην τεκμηριωμένη αξιοποίηση των δεδομένων για τη διαμόρφωση ενός ανταγωνιστικού και ασφαλούς αγωνιστικού περιβάλλοντος.

Λέξεις κλειδιά: Ανάλυση δεδομένων, αθλητισμός, τεχνολογία, απόδοση αθλητών, μηχανική μάθηση, πρόληψη τραυματισμών, στατιστικά & βιομετρικά δεδομένα.

ABSTRACT

This thesis explores the significance and practical applications of data analysis in the field of sports, highlighting the growing role of technology and science in enhancing athletic performance, strategy, and health management. It begins by presenting core concepts of data analysis, the types of data collected (player statistics, fitness metrics, biometric and psychological data), and the main sources used (wearables, sensors, video footage, lab testing). The study proceeds to examine data collection, preprocessing, and statistical analysis using modern tools such as Excel, Python, R, and SQL. Special focus is given to the use of machine learning algorithms for performance prediction, injury prevention, and strategic planning. Through real-world case studies, the research demonstrates that data analysis is not merely a supportive tool but an essential pillar of modern sports science and management. Ultimately, the study concludes that 21st-century sports heavily rely on informed, data-driven decision-making to build a competitive and safe athletic environment.

Keywords: Data analysis, sports, technology, athlete performance, machine learning, injury prevention, statistical & biometric data.

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

1.1 Ορισμός και σημασία της ανάλυσης δεδομένων στον αθλητισμό

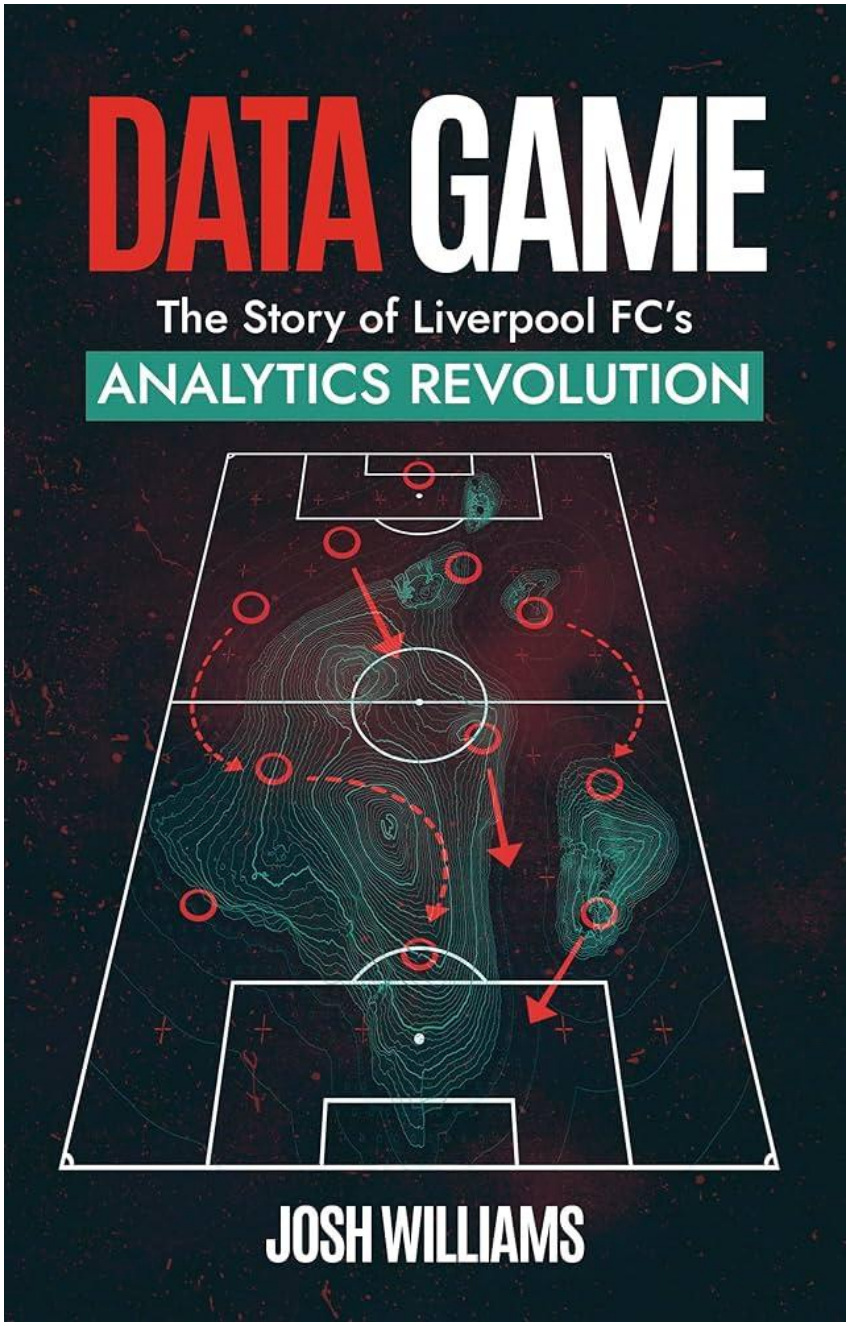
Η ανάλυση δεδομένων στον αθλητισμό αναφέρεται στη διαδικασία συλλογής, επεξεργασίας και ερμηνείας πληροφοριών σχετικών με την απόδοση αθλητών, ομάδων, αγώνων ή φυσιολογικών παραμέτρων. Οι πληροφορίες αυτές μπορεί να προέρχονται από στατιστικά αγώνων, συσκευές καταγραφής φυσικής κατάστασης, βιομετρικά δεδομένα, ακόμη και από βίντεο ή αισθητήρες GPS.

Η σημασία της ανάλυσης δεδομένων στον σύγχρονο αθλητισμό είναι τεράστια. Παρέχει πολύτιμα εργαλεία για τη λήψη αποφάσεων, τη βελτιστοποίηση της απόδοσης, την πρόληψη τραυματισμών, καθώς και για τη στρατηγική σχεδίαση τόσο σε ατομικό όσο και σε ομαδικό επίπεδο. Η τεχνολογία επιτρέπει στους προπονητές και τους επιστήμονες του αθλητισμού να βασίζονται σε αντικειμενικά δεδομένα και όχι μόνο σε εμπειρική γνώση ή παρατήρηση.

Από τις μεγάλες και επιτυχημένες ποδοσφαιρικές ομάδες και τις ομάδες του NBA, έως και ομάδες χαμηλότερου βεληνεκούς όπως σε πρωταθλήματα Σουηδίας, ακόμη και σε ιδιωτικούς συλλόγους που έχουν ως στόχο τη φυσική κατάσταση των αθλητών, οι αποφάσεις δεν βασίζονται πλέον μόνο στην εμπειρία και την παρατήρηση, αλλά και σε αντικειμενικά μετρήσιμα στοιχεία.

Για παράδειγμα:

- Η Λίβερπουλ στην Premier League αξιοποιεί προηγμένα συστήματα ανάλυσης (όπως το StatsBomb) για να αξιολογήσει την τακτική των αντιπάλων και την απόδοση των παικτών σε πραγματικό χρόνο.



Εικόνα 1: Το βιβλίο του Josh Williams για το data game της Liverpool FC

- Ομάδες του NBA χρησιμοποιούν αισθητήρες για να παρακολουθούν τη βιομηχανική κίνηση των παικτών τους, μειώνοντας τον κίνδυνο τραυματισμών.
- Η Formula 1 αποτελεί ίσως το πιο χαρακτηριστικό παράδειγμα όπου τεράστιος όγκος δεδομένων (δεκάδες GB ανά αγώνα) αναλύεται συνεχώς για την απόδοση του μονοθεσίου και του οδηγού.

- Ατομικά αθλήματα όπως το τρέξιμο ή η κολύμβηση ενσωματώνουν αναλυτικά dashboards για κάθε προπόνηση, τα οποία βοηθούν προπονητές και αθλητές να προσαρμόζουν τα πλάνα προπόνησης με βάση την πρόοδο ή την κόπωση.

Η εφαρμογή της ανάλυσης δεδομένων στον αθλητισμό δεν περιορίζεται μόνο σε επίπεδο αγωνιστικής απόδοσης, αλλά έχει επεκταθεί και στη διαχείριση αθλητών, τη διατροφή, τη διαχείριση κόπωσης, αλλά και σε οικονομικά και εμπορικά θέματα, όπως η στρατηγική μεταγραφών ή η ανάλυση της εμπειρίας των φιλάθλων.

Με την τεχνολογία να εξελίσσεται ραγδαία και τη διαθεσιμότητα μεγάλων ποσοτήτων δεδομένων (big data), η ανάλυση δεδομένων προσφέρει νέες δυνατότητες για προγνωστικά μοντέλα, στοχευμένη προπόνηση, και εξατομικευμένα πλάνα αποκατάστασης. Το αποτέλεσμα είναι μια πιο επιστημονικά τεκμηριωμένη, εξατομικευμένη και στρατηγική προσέγγιση του αθλητισμού σε όλα τα επίπεδα – από τον επαγγελματισμό έως τον μαζικό αθλητισμό και τη φυσική αγωγή.

1.2 **Στόχοι της εργασίας**

Η παρούσα εργασία έχει ως κύριο στόχο τη διερεύνηση του τρόπου με τον οποίο η ανάλυση δεδομένων μπορεί να εφαρμοστεί στον χώρο του αθλητισμού και να συμβάλει στη βελτίωση της απόδοσης και της στρατηγικής.

Ειδικότερα, επιδιώκεται:

- Η κατανόηση των βασικών εννοιών που σχετίζονται με την ανάλυση δεδομένων.
- Η παρουσίαση των τύπων δεδομένων που συλλέγονται στον αθλητισμό.
- Η παρουσίαση εργαλείων και μεθόδων που χρησιμοποιούνται για την ανάλυση αυτών των δεδομένων.
- Η εφαρμογή της θεωρίας σε πρακτικά παραδείγματα (case studies), όπως η αξιολόγηση της απόδοσης αθλητών, η πρόβλεψη αποτελεσμάτων και η πρόληψη τραυματισμών.

- Η ανάδειξη της σημασίας της τεχνολογίας και της επιστημονικής προσέγγισης στον αθλητισμό του 21ου αιώνα.

1.3 Μεθοδολογία και εργαλεία που θα χρησιμοποιηθούν

Η μεθοδολογία της εργασίας περιλαμβάνει βιβλιογραφική ανασκόπηση, παρουσίαση παραδειγμάτων ανάλυσης δεδομένων, και ανάλυση περιπτώσεων (case studies). Παράλληλα, γίνεται αναφορά σε σύγχρονα εργαλεία και τεχνολογίες που χρησιμοποιούνται στον χώρο του αθλητισμού για τη συλλογή και επεξεργασία δεδομένων.

1. Συλλογή Δεδομένων

Η πρώτη φάση αφορά την αναζήτηση και επιλογή κατάλληλων **πηγών δεδομένων**, οι οποίες μπορεί να περιλαμβάνουν:

- Δημόσιες βάσεις δεδομένων με στατιστικά αγώνων (π.χ. Kaggle, UEFA, NBA Stats).
- Ανοιχτά δεδομένα από εφαρμογές και τεχνολογικές πλατφόρμες (π.χ. Strava, WHOOP, Garmin Connect).
- Επιστημονική βιβλιογραφία και άρθρα που περιλαμβάνουν case studies.
- Αρχεία Excel ή CSV από μελέτες, αναλύσεις ή άλλες ερευνητικές πηγές.

2. Επεξεργασία και Καθαρισμός Δεδομένων

Τα δεδομένα που συλλέγονται συχνά περιέχουν ελλείψεις, σφάλματα ή πλεονάζουσες πληροφορίες. Συνεπώς, εφαρμόζονται τεχνικές όπως:

- Ανίχνευση και διόρθωση ελλιπών ή λανθασμένων τιμών.
- Κανονικοποίηση μονάδων (π.χ. λεπτά σε ώρες, kg σε gr).
- Αναδιοργάνωση δεδομένων σε κατάλληλη μορφή για ανάλυση.
- Αφαίρεση πλεοναστικών ή άσχετων πεδίων.

3. Ανάλυση Δεδομένων

Η ανάλυση πραγματοποιείται με τη χρήση τόσο βασικών στατιστικών μεθόδων, όσο και πιο προηγμένων τεχνικών, ανάλογα με το είδος των δεδομένων και το στόχο κάθε παραδείγματος. Ενδεικτικές τεχνικές:

- Περιγραφική στατιστική (μέσοι όροι, διακυμάνσεις, τυπικές αποκλίσεις).
- Γραφήματα τάσεων και σχέσεων (line plots, scatter plots, heatmaps).
- Ανάλυση παλινδρόμησης για πρόβλεψη αποτελεσμάτων.
- Συγκριτική ανάλυση μεταξύ ομάδων, περιόδων ή τύπων αθλητών.
- Εισαγωγή στην εφαρμογή αλγορίθμων μηχανικής μάθησης (machine learning) για την κατηγοριοποίηση ή πρόβλεψη δεδομένων.

4. Εργαλεία και Τεχνολογίες

Για την ανάλυση, θα χρησιμοποιηθούν δημοφιλή εργαλεία, εύκολα προσβάσιμα τόσο για ακαδημαϊκή όσο και για επαγγελματική χρήση:

- Microsoft Excel: Για βασική επεξεργασία, πίνακες, στατιστική και γραφήματα.
- Python: Γλώσσα προγραμματισμού με ισχυρές βιβλιοθήκες ανάλυσης όπως:
    - pandas (για διαχείριση δεδομένων),
    - matplotlib & seaborn (για γραφική απεικόνιση),
    - scikit-learn (για βασική μηχανική μάθηση).
- R: Εναλλακτική γλώσσα που ειδικεύεται στη στατιστική ανάλυση.
- SQL: Για τη διαχείριση και εξαγωγή δεδομένων από βάσεις δεδομένων (σε πιο προχωρημένο επίπεδο).
- Wearable technology platforms: Αναφορά σε τεχνολογίες όπως GPS trackers, HR monitors, και αθλητικά wearables.

5. Μελέτες Περίπτωσης (Case Studies)

Η εργασία θα εμπλουτιστεί με **παραδείγματα πραγματικών εφαρμογών**, όπου αναλύονται συγκεκριμένα σενάρια όπως:

- Ανάλυση της απόδοσης ποδοσφαιρικής ομάδας σε μια σεζόν.
- Χρήση δεδομένων για την πρόβλεψη της κόπωσης ή της πιθανότητας τραυματισμού.
- Σύγκριση αγωνιστικών χαρακτηριστικών μεταξύ δύο παικτών.

Η συνολική μεθοδολογική προσέγγιση στηρίζεται στην ιδέα ότι η ανάλυση δεδομένων στον αθλητισμό είναι ένα εργαλείο λήψης αποφάσεων, όχι απλώς μία θεωρητική διαδικασία. Γι' αυτό, δίνεται έμφαση στη σύνδεση της θεωρίας με την πράξη, και στην παρουσίαση χειροπιαστών εφαρμογών, κατανοητών ακόμη και χωρίς προηγούμενη εμπειρία σε στατιστική ή προγραμματισμό.

1.4 Δομή της εργασίας

Η εργασία είναι οργανωμένη ως εξής:

- **Κεφάλαιο 2**: Παρουσιάζει το θεωρητικό υπόβαθρο της ανάλυσης δεδομένων και τη σημασία της στον αθλητισμό.
- **Κεφάλαιο 3**: Αναλύει τις μεθόδους συλλογής, επεξεργασίας και ανάλυσης των δεδομένων, καθώς και τα εργαλεία που χρησιμοποιούνται.
- **Κεφάλαιο 4**: Παρουσιάζει μελέτες περίπτωσης (case studies) για την εφαρμογή της ανάλυσης δεδομένων στον αθλητισμό.
- **Κεφάλαιο 5**: Περιλαμβάνει τη συζήτηση, τα συμπεράσματα και τις προτάσεις για μελλοντική έρευνα και εφαρμογή.

ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 Βασικές έννοιες ανάλυσης δεδομένων

Ο αθλητισμός αποτελεί διαχρονικά έναν γόνιμο χώρο εφαρμογής της στατιστικής και της ανάλυσης δεδομένων, είτε για την πρόβλεψη αποτελεσμάτων, είτε για την αξιολόγηση παικτών και στρατηγικών κατά τη διάρκεια των αγώνων. Ειδικότερα στη σύγχρονη εποχή, με την εξέλιξη των τεχνολογιών καταγραφής και την ευρεία διάθεση δεδομένων, η ανάγκη για επιστημονική επεξεργασία και ανάλυση των αθλητικών πληροφοριών έχει γίνει πιο έντονη από ποτέ.

Χαρακτηριστικό παράδειγμα αποτελεί η περίπτωση της ομάδας μπέιζμπολ Oakland Athletics, η οποία βασίστηκε στην εμπειρική ανάλυση δεδομένων για να δημιουργήσει έναν ανταγωνιστικό πυρήνα παικτών, γεγονός που έγινε ευρέως γνωστό μέσω της ταινίας *Moneyball*. Σήμερα, στα μεγάλα πρωταθλήματα, ομάδες ειδικών αναλύουν εξονυχιστικά δεδομένα για να υποστηρίξουν στρατηγικές επιλογές, να προλάβουν τραυματισμούς, να βελτιώσουν την απόδοση των αθλητών και να ενισχύσουν τη θεαματικότητα του αθλήματος. Αυτή η τάση δεν αφορά μόνο τις επαγγελματικές ομάδες, αλλά αγγίζει και ακαδημαϊκούς, φιλάθλους και οργανισμούς που εμπλέκονται στον σχεδιασμό και την προβολή των αθλητικών διοργανώσεων.

Μέσα σε αυτό το πλαίσιο, η ανάλυση δεδομένων αναδεικνύεται σε απαραίτητο εργαλείο της αθλητικής επιστήμης, καθώς περιλαμβάνει τη συλλογή, την επεξεργασία, την ερμηνεία και την αξιοποίηση πληροφοριών για την κατανόηση σύνθετων φαινομένων. Μεθοδολογίες που βασίζονται στα μαθηματικά, τη στατιστική και τη μηχανική μάθηση βοηθούν ομάδες, προπονητές και φιλάθλους να λαμβάνουν τεκμηριωμένες αποφάσεις, να βελτιώνουν τις επιδόσεις τους ή ακόμα και να απολαμβάνουν μια πιο εμπλουτισμένη εμπειρία θέασης.

Βασικές έννοιες που σχετίζονται με την ανάλυση δεδομένων:

- **Δεδομένα (Data)**: Πληροφορίες σε μορφή αριθμών, κειμένων ή μετρήσεων.
- **Μεταβλητές**: Ποσότητες που καταγράφονται, π.χ. καρδιακός ρυθμός, απόσταση, διάρκεια.

- **Στατιστική ανάλυση**: Εργαλεία και μέθοδοι για να κατανοούμε και να συγκρίνουμε τα δεδομένα.
- **Οπτικοποίηση δεδομένων**: Δημιουργία γραφημάτων και πινάκων για καλύτερη ερμηνεία.
- **Predictive analytics**: Τεχνικές που χρησιμοποιούνται για την πρόβλεψη μελλοντικών γεγονότων.

Η σωστή κατανόηση αυτών των εννοιών είναι θεμέλιο για την εφαρμογή της ανάλυσης δεδομένων στην πράξη, είτε πρόκειται για προπονητική αξιολόγηση, είτε για αγωνιστική στρατηγική.

Δεδομένα (Data):

Τα δεδομένα αποτελούν το θεμέλιο κάθε διαδικασίας ανάλυσης. Πρόκειται για πληροφορίες που έχουν συλλεχθεί με σκοπό την επεξεργασία, την κατανόηση και τη λήψη αποφάσεων. Στο πλαίσιο του αθλητισμού, τα δεδομένα μπορεί να περιλαμβάνουν αριθμητικές τιμές όπως οι επιδόσεις ενός αθλητή, ο αριθμός των επαναλήψεων σε μια άσκηση, ή ακόμη και πιο σύνθετες πληροφορίες, όπως βιομετρικά χαρακτηριστικά και στατιστικά αγώνων. Τα δεδομένα μπορεί να είναι δομημένα (π.χ. πίνακες Excel με μετρήσεις) ή μη δομημένα (π.χ. βίντεο ή χειρόγραφες σημειώσεις).

Μεταβλητές:

Οι μεταβλητές είναι συγκεκριμένα στοιχεία ή χαρακτηριστικά που μετρούνται και καταγράφονται μέσα στα δεδομένα. Στην ανάλυση δεδομένων, κάθε μεταβλητή αντιπροσωπεύει ένα είδος πληροφορίας, όπως για παράδειγμα ο καρδιακός ρυθμός, η διάρκεια μιας προπόνησης, η απόσταση που διανύθηκε ή το ποσοστό λίπους στο σώμα. Οι μεταβλητές μπορεί να είναι ποσοτικές (αριθμητικές) ή ποιοτικές (κατηγορικές), και η σωστή επιλογή και καταγραφή τους είναι καθοριστικής σημασίας για την ακρίβεια και την αξιοπιστία της ανάλυσης.

Στατιστική ανάλυση:

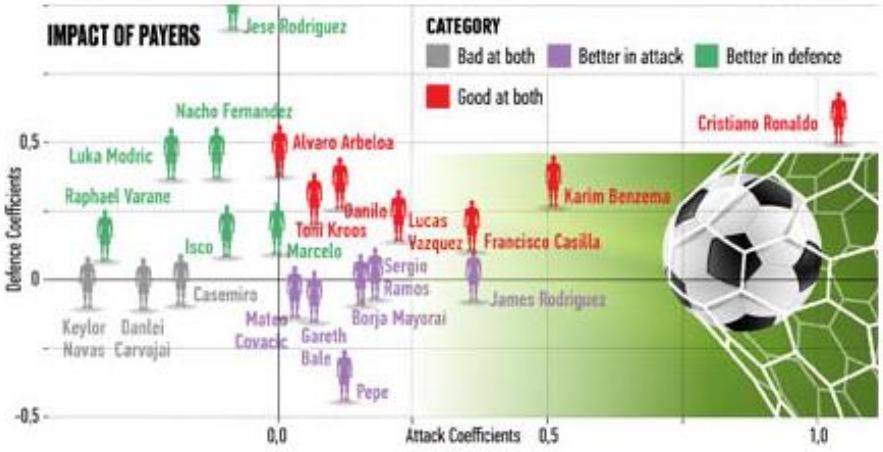
Η στατιστική ανάλυση περιλαμβάνει ένα σύνολο εργαλείων και μεθόδων που μας βοηθούν να εξετάσουμε και να ερμηνεύσουμε τα δεδομένα. Μέσω αυτής, μπορούμε να εντοπίσουμε τάσεις, σχέσεις και αποκλίσεις, καθώς και να συγκρίνουμε διαφορετικά σύνολα δεδομένων. Για παράδειγμα, μπορούμε να μελετήσουμε πώς επηρεάζει ο χρόνος ύπνου την αθλητική απόδοση ή να συγκρίνουμε τις επιδόσεις δύο ομάδων σε διαφορετικές χρονικές περιόδους. Η στατιστική είναι απαραίτητη για την εξαγωγή ουσιαστικών συμπερασμάτων από μεγάλα και πολύπλοκα δεδομένα.

Οπτικοποίηση δεδομένων:

Η οπτικοποίηση δεδομένων είναι η διαδικασία παρουσίασης των πληροφοριών με τη χρήση γραφημάτων, διαγραμμάτων, πινάκων ή και διαδραστικών εργαλείων. Μέσα από την οπτική αναπαράσταση, γίνεται πιο εύκολη η κατανόηση των αριθμών και των σχέσεων μεταξύ των μεταβλητών. Ένα γράφημα με την εξέλιξη των επιδόσεων ενός αθλητή ανά εβδομάδα είναι πιο κατανοητό από έναν απλό πίνακα αριθμών. Η σωστή οπτικοποίηση διευκολύνει την επικοινωνία των αποτελεσμάτων και τη λήψη αποφάσεων από προπονητές, αθλητές και επιστήμονες του αθλητισμού.

Predictive analytics (Αναλυτική πρόβλεψη):

Η predictive analytics αναφέρεται στη χρήση προηγμένων τεχνικών, όπως στατιστικά μοντέλα και μηχανική μάθηση, για την πρόβλεψη μελλοντικών γεγονότων με βάση υπάρχοντα δεδομένα. Στον αθλητισμό, μπορεί να χρησιμοποιηθεί για την πρόβλεψη τραυματισμών, την εκτίμηση της μελλοντικής απόδοσης ενός αθλητή ή την ανάλυση στρατηγικών αντιπάλων. Πρόκειται για ένα ισχυρό εργαλείο που δίνει τη δυνατότητα στους επαγγελματίες να προετοιμαστούν καλύτερα και να λάβουν πιο τεκμηριωμένες αποφάσεις.



Εικόνα 2: Αποτέλεσμα γραφήματος από τις παραπάνω μεταβλητές στους παίκτες της Ρεάλ Μαδρίτης

2.2 Είδη αθλητικών δεδομένων

Η ανάλυση δεδομένων στον χώρο του αθλητισμού βασίζεται σε ποικίλους τύπους πληροφοριών, που συλλέγονται κατά τη διάρκεια της προπόνησης, των αγώνων και της καθημερινής φυσικής δραστηριότητας των αθλητών. Αυτά τα δεδομένα διαχωρίζονται σε κατηγορίες όπως στατιστικά παικτών, δεδομένα αγώνων, φυσικής κατάστασης, βιομετρικά στοιχεία και πολλά ακόμη. Καθεμιά από αυτές τις κατηγορίες αποτελεί πηγή μεταβλητών και δεδομένων που μπορούν να υποστούν στατιστική ανάλυση, να οπτικοποιηθούν και να αξιοποιηθούν ακόμα και για πρόβλεψη μελλοντικών επιδόσεων μέσω predictive analytics.

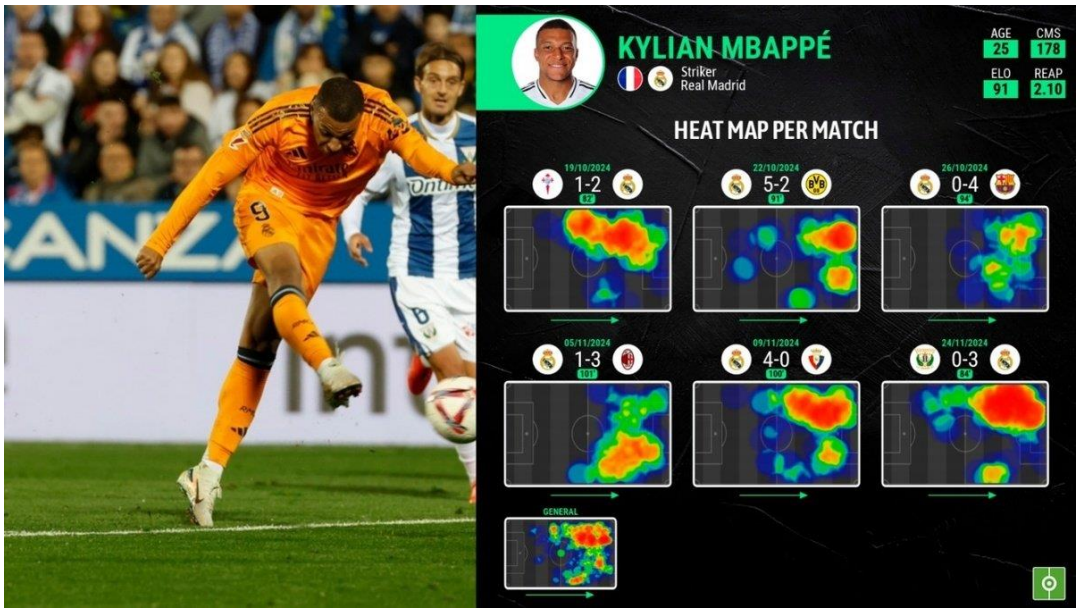
1. Στατιστικά παικτών (Player statistics):

Αυτά τα δεδομένα αφορούν μεμονωμένους αθλητές και περιλαμβάνουν αριθμητικές πληροφορίες όπως πόντοι, ασίστ, επιτυχημένες πάσες, καρδιοαναπνευστικές μετρήσεις, χρόνος συμμετοχής, κ.λπ. Αποτελούν ποσοτικά δεδομένα, όπου κάθε τιμή συνιστά μια μεταβλητή που καταγράφεται συστηματικά. Μέσω στατιστικής ανάλυσης, μπορούμε να εντοπίσουμε την απόδοση ενός παίκτη σε σχέση με προηγούμενες εμφανίσεις ή με τη μέση απόδοση της ομάδας. Η οπτικοποίηση αυτών των στατιστικών (π.χ. με ραβδογράμματα ή ραντάρ γραφήματα) βοηθά προπονητές και παίκτες να κατανοήσουν δυνατά και αδύναμα σημεία. Επιπλέον, χρησιμοποιώντας

predictive analytics, μπορεί να προβλεφθεί η απόδοση ενός αθλητή σε επόμενους αγώνες, βάσει της τρέχουσας φόρμας και ιστορικών στοιχείων.

2. Δεδομένα αγώνων (Match statistics):

Αυτά περιλαμβάνουν όλες τις πληροφορίες που αφορούν την έκβαση ενός αγώνα, τη ροή, την τακτική και τη συμμετοχή της κάθε ομάδας. Παραδείγματα είναι το ποσοστό κατοχής, τα σουτ στην εστία, οι αμυντικές ενέργειες, τα λάθη, η απόσταση που διανύθηκε από κάθε παίκτη κ.ά. Πρόκειται για σύνθετα σύνολα δεδομένων, με πολλές μεταβλητές που εξετάζονται σε σχέση μεταξύ τους. Η στατιστική ανάλυση μπορεί να αποκαλύψει τις τακτικές που οδήγησαν στη νίκη ή την ήττα, ενώ η οπτικοποίηση σε διαγράμματα θερμότητας (heatmaps) ή χρονοδιαγράμματα βοηθά στην καλύτερη κατανόηση της αγωνιστικής δραστηριότητας. Με την αξιοποίηση predictive models, είναι δυνατή η πρόβλεψη του αποτελέσματος μελλοντικών αγώνων ή η εκτίμηση πιθανότητας τραυματισμού λόγω κόπωσης.



Εικόνα 3: Το heatmap του ποδοσφαιριστή Kylian Mbappe σε ένα εύρος αγώνων

3. Δεδομένα φυσικής κατάστασης (Fitness & performance data): Αυτά τα δεδομένα προέρχονται κυρίως από αισθητήρες, wearable συσκευές και εργομετρικά τεστ. Περιλαμβάνουν τον καρδιακό ρυθμό, τα επίπεδα γαλακτικού οξέος, τη VO_2 max, την ταχύτητα, την επιτάχυνση, τον χρόνο αντίδρασης και άλλα βιομετρικά χαρακτηριστικά. Πρόκειται για δεδομένα υψηλής ακρίβειας που

σχετίζονται άμεσα με τη φυσική κατάσταση και το φορτίο προπόνησης. Οι μεταβλητές εδώ είναι πολλές και ευαίσθητες, και η στατιστική ανάλυσή τους βοηθά στον προγραμματισμό της προπόνησης, στην αποφυγή υπερκόπωσης και στην αποκατάσταση. Με τη σωστή οπτικοποίηση (π.χ. διαγράμματα κόπωσης ή φόρτου), οι ειδικοί μπορούν να προσαρμόσουν άμεσα το πλάνο. Μέσω predictive analytics, προβλέπεται η βελτίωση της φυσικής κατάστασης ή ακόμη και ο κίνδυνος τραυματισμού.

4. Βιομετρικά και ιατρικά δεδομένα (Biometric & medical data): Περιλαμβάνουν μετρήσεις όπως βάρος, ύψος, σύσταση σώματος (λιπώδης/μυϊκή μάζα), αλλά και ιατρικά ιστορικά, τραυματισμούς, χρόνιες παθήσεις. Είναι δεδομένα κρίσιμα για την εξατομικευμένη προσέγγιση στην προπόνηση και στη φροντίδα των αθλητών. Αυτές οι μεταβλητές πρέπει να αντιμετωπίζονται με υψηλό βαθμό εμπιστευτικότητας, αλλά και ακρίβειας. Η στατιστική ανάλυση μπορεί να αποκαλύψει συσχετίσεις (π.χ. σύνδεση υπέρβαρου με αυξημένο κίνδυνο τραυματισμού), ενώ η οπτικοποίηση μπορεί να βοηθήσει στην παρακολούθηση της προόδου υγείας ενός αθλητή. Μέσω προγνωστικών μοντέλων, είναι δυνατό να εκτιμηθεί ο χρόνος αποκατάστασης ή η πιθανότητα υποτροπής ενός τραυματισμού.

5. Ψυχολογικά και γνωστικά δεδομένα (Psychological & cognitive data): Τα δεδομένα αυτά σχετίζονται με την ψυχολογική κατάσταση, το στρες, το επίπεδο συγκέντρωσης και τη νοητική κόπωση των αθλητών. Παρόλο που είναι πιο δύσκολο να μετρηθούν αντικειμενικά, μπορούν να συλλεχθούν μέσω ερωτηματολογίων, τεστ γνωστικών ικανοτήτων ή ακόμα και μέσω βιοαισθητήρων. Είναι ποιοτικά και ποσοτικά δεδομένα ταυτόχρονα, που εμπλουτίζουν την ανάλυση και επιτρέπουν πιο σφαιρική αξιολόγηση της κατάστασης του αθλητή. Η ανάλυση αυτών των δεδομένων, σε συνδυασμό με τα υπόλοιπα, προσφέρει ένα πλήρες προφίλ του αθλητή, ενώ με predictive analytics μπορεί να εκτιμηθεί η πνευματική ετοιμότητα για έναν κρίσιμο αγώνα.

2.3. Πηγές συλλογής αθλητικών δεδομένων

Η ανάλυση δεδομένων στον αθλητισμό ξεκινά με τη συλλογή αξιόπιστων πληροφοριών από διάφορες πηγές. Οι πηγές αυτές ποικίλλουν ανάλογα με το είδος του αθλητισμού, το επίπεδο του αθλητή (επαγγελματικό ή ερασιτεχνικό), και τους

στόχους της ανάλυσης (π.χ. βελτίωση απόδοσης, πρόληψη τραυματισμών, στρατηγική ομάδας). Η ποιότητα και η ακρίβεια των δεδομένων καθορίζουν σε μεγάλο βαθμό την αξιοπιστία των αποτελεσμάτων της ανάλυσης.

Οι συσκευές που χρησιμοποιούνται για την άντληση δεδομένων από αθλητές είναι οι παρακάτω:

Wearable συσκευές και αισθητήρες (π.χ. GPS trackers, καρδιοσυχνόμετρα):

Οι φορητές τεχνολογίες αποτελούν σήμερα μια από τις πιο διαδεδομένες πηγές συλλογής δεδομένων φυσικής κατάστασης. Οι συσκευές αυτές καταγράφουν ζωτικά στοιχεία όπως ο καρδιακός ρυθμός, η ταχύτητα, η απόσταση, οι επιταχύνσεις, οι αλλαγές κατεύθυνσης και η συνολική απόδοση κατά τη διάρκεια προπονήσεων και αγώνων. Παρέχουν συνεχές ρεύμα ποσοτικών δεδομένων, τα οποία είναι κρίσιμα για στατιστική ανάλυση, παρακολούθηση κόπωσης, και πρόγνωση φυσικής επιβάρυνσης.



Εικόνα 4: GPS Tracker σε αθλητές (STATSports.com)

Βίντεο ανάλυση (Video analysis):

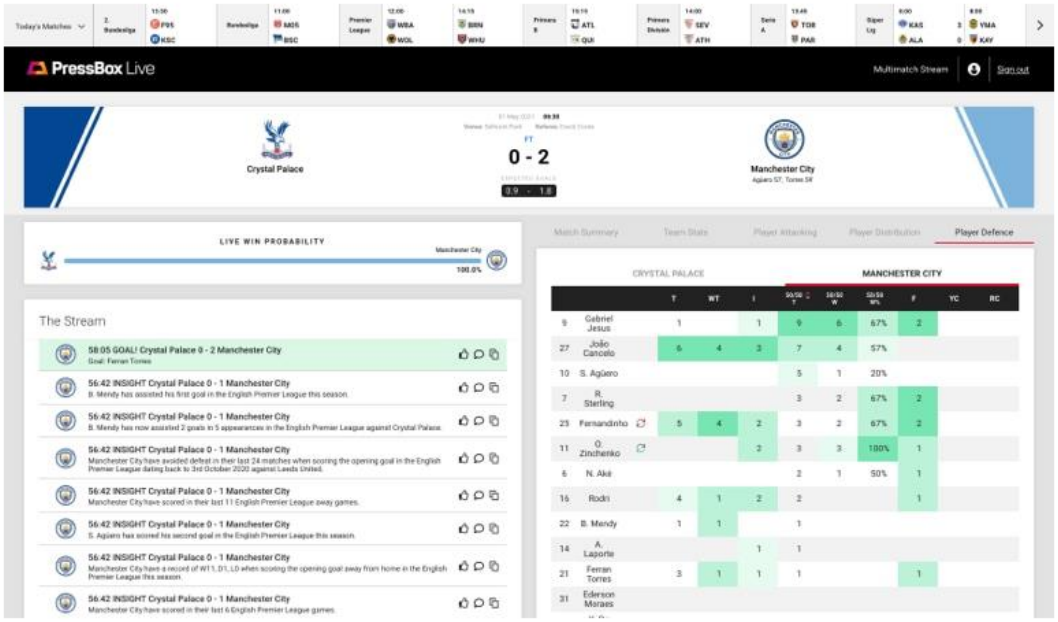
Η χρήση βίντεο σε συνδυασμό με εξειδικευμένο λογισμικό επιτρέπει την ποιοτική και ποσοτική αξιολόγηση της απόδοσης παικτών και ομάδων. Αναλύονται κινήσεις, τακτικές, τοποθετήσεις και λάθη, ενώ συνδέεται και με οπτικοποίηση δεδομένων, π.χ. μέσω heatmaps. Το video analysis παρέχει πληροφορίες που δεν καταγράφονται από αισθητήρες, ενισχύοντας τη συνολική εικόνα ενός αγώνα ή μιας προπόνησης.



Εικόνα 5: Ανάλυση του αγώνα μέσω Video Analysis

Στατιστικές βάσεις δεδομένων αγώνων και παικτών:

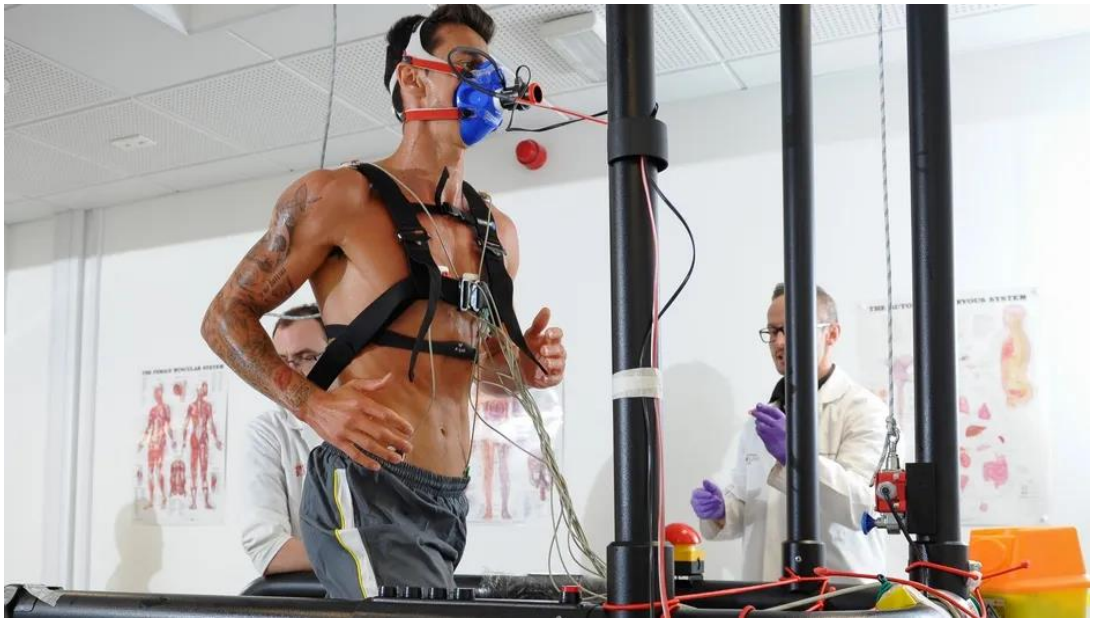
Διάφορες επίσημες πλατφόρμες (π.χ. Opta, InStat, StatsPerform, UEFA database) προσφέρουν ολοκληρωμένα στατιστικά δεδομένα για επαγγελματικές ομάδες και αθλητές. Αυτές οι δομημένες βάσεις δεδομένων παρέχουν άμεσα στατιστικά παικτών και αγώνων, επιτρέποντας σύγκριση επιδόσεων μεταξύ παικτών ή ομάδων και ανάλυση ιστορικών τάσεων. Αποτελούν θεμέλιο για predictive analytics, καθώς βασίζονται σε τεράστιους όγκους δεδομένων (big data).



Εικόνα 6: Opta Live Stats Performance (Opta.com)

Εργομετρικά και ιατρικά τεστ:

Η εργομετρία περιλαμβάνει δοκιμές όπως η μέτρηση VO_2 max, τεστ ευλυγισίας, εκρηκτικότητας και δύναμης. Αυτά τα δεδομένα συλλέγονται σε ειδικά εργαστήρια ή με επαγγελματικό εξοπλισμό και είναι εξαιρετικά χρήσιμα για τον προγραμματισμό προπόνησης και τη βελτίωση φυσικής κατάστασης. Επίσης, ιατρικές εξετάσεις (π.χ. αιματολογικές, απεικονιστικές) προσφέρουν δεδομένα για την υγεία του αθλητή, που είναι απαραίτητα για την πρόληψη τραυματισμών και την προγνωστική παρακολούθηση της φυσικής φθοράς.



Εικόνα 7: Αθλητής σε ιατρικό τεστ ώστε να αντληθούν δεδομένα

Ψυχολογικά ερωτηματολόγια και τεστ αξιολόγησης:

Εργαλεία όπως ερωτηματολόγια διάθεσης, στρες, επαγγελματικής εξουθένωσης, αλλά και τεστ αντίδρασης ή συγκέντρωσης, συλλέγουν ψυχολογικά και γνωστικά δεδομένα. Παρότι πιο δύσκολο να ποσοτικοποιηθούν, είναι κρίσιμα για την ολιστική ανάλυση της κατάστασης του αθλητή και συχνά χρησιμοποιούνται συμπληρωματικά με τα βιολογικά δεδομένα.

Χειροκίνητη καταγραφή δεδομένων (π.χ. μέσω προπονητών ή αναλυτών):

Σε χαμηλότερα επίπεδα αθλητισμού, όπου δεν υπάρχουν προηγμένες τεχνολογίες, τα δεδομένα καταγράφονται από τους προπονητές ή τους παρατηρητές, π.χ. με καταγραφή λαθών, χρονισμού ή παρατηρήσεις σχετικά με την κόπωση ή την τεχνική. Αν και λιγότερο ακριβή, αποτελούν σημαντική ποιοτική πληροφορία και συνήθως συνδυάζονται με τα υπόλοιπα δεδομένα για καλύτερη κατανόηση.



Εικόνα 8: Ο προπονητής Λουίς Φαν Χάαλ σημειώνει κατά τη διάρκεια ενός αγώνα

Οι πηγές συλλογής δεδομένων στον αθλητισμό ποικίλουν από τεχνολογικά εργαλεία υψηλής ακρίβειας μέχρι παραδοσιακές μεθόδους καταγραφής. Η κατανόηση των χαρακτηριστικών κάθε πηγής επιτρέπει τη βελτιστοποίηση της συλλογής και ανάλυσης δεδομένων, οδηγώντας σε πιο τεκμηριωμένες αποφάσεις και στρατηγικές.

2.4 Εφαρμογή της Ανάλυσης Δεδομένων στον Αθλητισμό

Στο υποκεφάλαιο 2.3, αναδείχθηκε ο κρίσιμος ρόλος των σύγχρονων συστημάτων καταγραφής και ανάλυσης δεδομένων στον αθλητισμό. Μέσω αυτών των τεχνολογιών, καθίσταται εφικτή η συστηματική συλλογή πληροφοριών που αφορούν την απόδοση, τη φυσική κατάσταση, τις κινήσεις και τις στρατηγικές των αθλητών και των ομάδων. Ωστόσο, η απλή καταγραφή δεν επαρκεί. Η πραγματική αξία των δεδομένων αναδεικνύεται μέσα από τη σωστή ανάλυσή τους, η οποία οδηγεί σε ουσιαστικά συμπεράσματα και πρακτικές εφαρμογές. Ο Haefner (n.d.) επισημαίνει τη σημασία της παρακολούθησης βασικών δεικτών όπως τα ριμπάουντ και τα turnovers στην ανάλυση απόδοσης μπάσκετ.

Η ανάλυση δεδομένων στον αθλητισμό έχει εξελιχθεί σε ένα ισχυρό εργαλείο για την ενίσχυση της αθλητικής απόδοσης, τη λήψη τεκμηριωμένων αποφάσεων και τη βελτίωση όλων των πτυχών της αθλητικής δραστηριότητας – από την προπόνηση έως

τη διαχείριση τραυματισμών και την εμπειρία των θεατών. Στην παρούσα ενότητα παρουσιάζονται αναλυτικά οι βασικότερες εφαρμογές της.

1. Παρακολούθηση και Βελτίωση της Αθλητικής Απόδοσης

Μια από τις κύριες εφαρμογές της ανάλυσης δεδομένων είναι η παρακολούθηση της απόδοσης των αθλητών σε πραγματικό χρόνο και η μακροχρόνια καταγραφή της εξέλιξής τους. Μέσω συσκευών όπως GPS, επιταχυνσιόμετρα, καρδιοσυχνόμετρα και wearable τεχνολογία (π.χ. Polar, Catapult, Garmin), είναι δυνατή η μέτρηση σημαντικών μεταβλητών όπως (Stats Perform, n.d.):

- Καρδιακή συχνότητα και ζώνες έντασης
- Χρόνος αποκατάστασης
- Απόσταση που διανύθηκε
- Αριθμός σπριντ, αλλαγές κατεύθυνσης, επιταχύνσεις/επιβραδύνσεις
- Κατανάλωση ενέργειας

Αναλύοντας αυτά τα δεδομένα, οι προπονητές προσαρμόζουν την ένταση και τη διάρκεια των προπονήσεων ανάλογα με τις ανάγκες κάθε αθλητή, βελτιστοποιώντας έτσι τη φυσική του κατάσταση χωρίς να τον υπερφορτώνουν. Αυτό οδηγεί σε στοχευμένη προπόνηση και ατομικοποίηση του προπονητικού φορτίου, κάτι ιδιαίτερα σημαντικό στον υψηλό ανταγωνισμό.

2. Ανάλυση Αντιπάλων και Αγωνιστικής Στρατηγικής

Η στρατηγική προετοιμασία για έναν αγώνα βασίζεται ολοένα και περισσότερο στην ανάλυση δεδομένων αντιπάλων. Χρησιμοποιώντας λογισμικά όπως το Hudl, το Wyscout ή το Instat, αναλυτές αγώνων επεξεργάζονται στοιχεία όπως:

- Μοτίβα κίνησης και ανάπτυξης παιχνιδιού
- Συχνότητα πάσας ανά ζώνη γηπέδου
- Πιθανές αδυναμίες στην άμυνα ή στα στημένα
- Ποσοστά επιτυχημένων ενεργειών (π.χ. τάκλιν, κλεψίματα, σουτ εντός εστίας)

Με βάση αυτά, οι προπονητές χαράζουν προσαρμοσμένα πλάνα τακτικής, π.χ. επιλέγοντας πρέσινγκ σε συγκεκριμένες περιοχές ή προσαρμόζοντας τη διάταξη της ομάδας ώστε να εκμεταλλευτούν τακτικά κενά του αντιπάλου. Έτσι, η στατιστική ανάλυση μετατρέπεται σε αγωνιστικό πλεονέκτημα.

3. Πρόληψη Τραυματισμών και Αποκατάσταση

Η ανάλυση δεδομένων είναι καθοριστική και για την πρόληψη τραυματισμών, ένας τομέας που απασχολεί ιδιαίτερα τα επαγγελματικά σωματεία. Μέσω εργαλείων καταγραφής του προπονητικού φορτίου (internal και external load), γίνεται συνεχής παρακολούθηση της επιβάρυνσης που δέχεται ένας αθλητής. Η απότομη αύξηση της έντασης αποτελεί παράγοντα κινδύνου για μυϊκούς τραυματισμούς.

Με predictive analytics (προγνωστική ανάλυση), οι προπονητές μπορούν να εντοπίζουν red flags, δηλαδή ενδείξεις υπερκόπωσης ή ασύμμετρης καταπόνησης, και να προσαρμόζουν τα φορτία εγκαίρως. Επιπλέον, κατά τη φάση της αποκατάστασης, τα δεδομένα βοηθούν στην παρακολούθηση της προόδου και στην επαναφορά του αθλητή με ασφάλεια και σταδιακά, χωρίς να επιβαρυνθεί πρόωρα (Pezzana, 2021).

4. Στρατολόγηση Παικτών (Scouting)

Η ανάλυση δεδομένων έχει μεταμορφώσει τον τρόπο με τον οποίο οι ομάδες εντοπίζουν και επιλέγουν παίκτες. Πλέον, δεν βασίζονται μόνο στην οπτική παρατήρηση ή την εμπειρική αξιολόγηση, αλλά χρησιμοποιούν ποσοτικά κριτήρια για να εντοπίσουν ποδοσφαιριστές ή αθλητές με συγκεκριμένα χαρακτηριστικά.

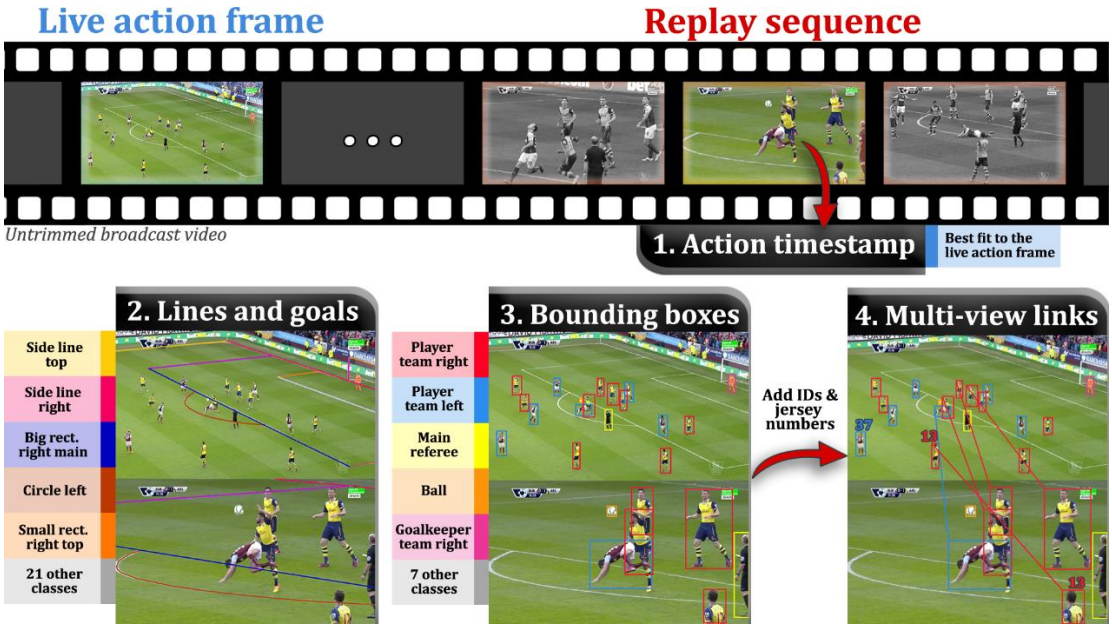
Στις βάσεις δεδομένων scouting, κάθε παίκτης αξιολογείται με βάση δεκάδες δείκτες απόδοσης (KPIs), όπως:

- Ποσοστά επιτυχημένων πασών
- Ανακτήσεις μπάλας ανά αγώνα
- Σουτ ανά προσπάθεια και ποσοστά ευστοχίας
- Συμμετοχή στη δημιουργία φάσεων

Έτσι, η στρατολόγηση γίνεται πιο αντικειμενική, στοχευμένη και αποτελεσματική, μειώνοντας τον κίνδυνο λανθασμένων μεταγραφών.



Εικόνα 9: Scouting εντός γηπέδου



Εικόνα 10: Scouting με video analysis – και data analytics εκτός γηπέδου

5. Εμπειρία Θεατών και Εμπορική Εκμετάλλευση

Η ανάλυση δεδομένων χρησιμοποιείται και για τη βελτίωση της εμπειρίας των φιλάθλων. Οι σύγχρονες μεταδόσεις αγώνων συνοδεύονται από real-time στατιστικά, θερμογραφικές απεικονίσεις κινήσεων, χάρτες πίεσης και διαδραστικά γραφήματα, κάνοντας την παρακολούθηση πιο ενδιαφέρουσα και διαδραστική.

Παράλληλα, οι σύλλογοι χρησιμοποιούν big data για στοχευμένο marketing, δυναμική τιμολόγηση εισιτηρίων και ανάλυση συμπεριφοράς φιλάθλων (π.χ. ποιοι αγοράζουν διαρκείας, πότε, σε ποιους αγώνες παραβρίσκονται). Αυτό ενισχύει τα έσοδα και τη σύνδεση με το κοινό.



Εικόνα 11: Real-time data στο ημίχρονο ενός αγώνα μπάσκετ

Συμπεράσματα

Εν ολίγοις, Η ανάλυση δεδομένων έχει μετατραπεί σε αναπόσπαστο εργαλείο κάθε σοβαρής αθλητικής οντότητας. Από την ατομική προπόνηση έως τον στρατηγικό σχεδιασμό ολόκληρης σεζόν, και από τη στρατολόγηση παικτών έως την εμπορική αξιοποίηση, οι εφαρμογές της είναι πολυεπίπεδες και καθοριστικές για την επιτυχία. Η επιστημονική προσέγγιση που προσφέρει συμβάλλει καθοριστικά στη μεγιστοποίηση της απόδοσης και την ελαχιστοποίηση των κινδύνων, προάγοντας έναν πιο σύγχρονο, έξυπνο και αποτελεσματικό αθλητισμό. Δεν πρόκειται απλώς για ένα εργαλείο υποστήριξης, αλλά για έναν κομβικό πυλώνα λήψης αποφάσεων σε όλα τα επίπεδα: ατομικό, ομαδικό, τεχνικό, διοικητικό και εμπορικό. Η επίδρασή της αγγίζει όχι μόνο την απόδοση των αθλητών, αλλά και τον σχεδιασμό των προπονήσεων, την υγειονομική διαχείριση, την τεχνική προσέγγιση των αγώνων, καθώς και τη συνολική λειτουργία των συλλόγων και των ομοσπονδιών (Cai, Zhao & Wang, 2021).

Ακόμα και σε χαμηλότερα επίπεδα ή σε ερασιτεχνικά πρωταθλήματα, η εισαγωγή απλών μορφών καταγραφής (όπως εφαρμογές GPS, στατιστικά κ.λπ.) μπορεί να βελτιώσει θεαματικά την ποιότητα της προπόνησης και του αγωνιστικού σχεδιασμού.

Σε ένα ευρύτερο πλαίσιο, η ανάλυση δεδομένων ενισχύει την εξέλιξη του αθλητισμού ως πλέον μία επιστήμη, δημιουργώντας γέφυρες ανάμεσα στην τεχνολογία, τη φυσική αγωγή, τη βιολογία και τη στατιστική. Καθώς τα δεδομένα αυξάνονται και οι μέθοδοι ανάλυσης εξελίσσονται, το μέλλον του αθλητισμού διαμορφώνεται σε έναν χώρο όπου η πληροφορία είναι δύναμη, και η ικανότητα αξιοποίησής της αποτελεί συγκριτικό πλεονέκτημα.

Ο αθλητισμός του 21ου αιώνα δεν βασίζεται μόνο στο ταλέντο και την προσπάθεια, αλλά και στην ικανότητα κατανόησης και αξιοποίησης των δεδομένων. Όσοι φορείς, αθλητές και οργανισμοί ενσωματώνουν αυτή τη νέα πραγματικότητα, διαμορφώνουν τους αυριανούς πρωταγωνιστές. Όπως αναφέρει ο Hayhurst (2020), τα analytics έχουν αλλάξει ριζικά τον τρόπο με τον οποίο οι ομάδες NBA αξιολογούν τους παίκτες.

ΚΕΦΑΛΑΙΟ 3: ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΕΡΓΑΛΕΙΑ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ ΣΤΟΝ ΑΘΛΗΤΙΣΜΟ

3.1 Εισαγωγή – Συλλογή δεδομένων (πηγές, μέθοδοι, αξιοπιστία)

Η ανάλυση δεδομένων στον αθλητισμό απαιτεί μια συστηματική μεθοδολογική προσέγγιση, προκειμένου τα αποτελέσματα να είναι αξιόπιστα, χρήσιμα και εφαρμόσιμα στην πράξη. Από τη συλλογή των δεδομένων έως την τελική ανάλυση, κάθε στάδιο είναι κρίσιμο για τη λήψη σωστών αποφάσεων που επηρεάζουν τις επιδόσεις, την υγεία των αθλητών και τη στρατηγική των ομάδων. Στο παρόν κεφάλαιο περιγράφονται οι βασικές μεθοδολογίες και τα εργαλεία που χρησιμοποιούνται στην ανάλυση δεδομένων στον αθλητισμό, με στόχο την καλύτερη κατανόηση της διαδικασίας και των τεχνολογικών επιλογών που τη στηρίζουν. Η πλατφόρμα Hudl προσφέρει ολοκληρωμένες λύσεις για την καταγραφή και ανάλυση αθλητικών δεδομένων (Hudl Blog, 2016)

Η συλλογή δεδομένων αποτελεί το πρώτο και θεμελιώδες στάδιο στην αναλυτική διαδικασία. Οι πηγές μπορεί να είναι ποικίλες:

- **Πηγές**: wearables (ρολόγια, GPS, καρδιοσυχνόμετρα), συστήματα video-tracking, στατιστικές πλατφόρμες (π.χ. Instat, Hudl), ερωτηματολόγια, ιατρικά και προπονητικά αρχεία.
- **Μέθοδοι**: παρατήρηση, καταγραφή με αισθητήρες, αυτόματη εξαγωγή δεδομένων από λογισμικά ή APIs, χειροκίνητη καταγραφή από προπονητές ή αναλυτές.
- **Αξιοπιστία**: εξαρτάται από την ακρίβεια του εξοπλισμού, τη συχνότητα μέτρησης και τη σταθερότητα του περιβάλλοντος (π.χ. εξωτερικός χώρος με GPS). Η διασφάλιση της αξιοπιστίας είναι κρίσιμη ώστε τα συμπεράσματα να είναι έγκυρα.

Όπως παρουσιάστηκε στο Κεφάλαιο 2, η σύγχρονη αθλητική πραγματικότητα βασίζεται σε ένα μεγάλο εύρος δεδομένων, τα οποία προκύπτουν είτε από τεχνολογικά συστήματα καταγραφής (όπως wearables, GPS, αισθητήρες, βιντεοανάλυση) είτε από

στατιστικά παικτών και αγώνων. Τα δεδομένα αυτά αποτελούν τη βάση για οποιαδήποτε διαδικασία ανάλυσης, καθώς χωρίς ποιοτική και αξιόπιστη συλλογή δεν μπορεί να προκύψει χρήσιμο συμπέρασμα.

Στο παρόν υποκεφάλαιο δεν θα επαναλάβουμε τις τεχνικές λεπτομέρειες ή τα είδη των δεδομένων που ήδη αναλύθηκαν στο Κεφάλαιο 2 (2.2 και 2.3), αλλά θα εστιάσουμε στο μεθοδολογικό πλαίσιο της συλλογής, ως κρίσιμο πρώτο βήμα στην ευρύτερη διαδικασία της ανάλυσης δεδομένων.

Η συλλογή περιλαμβάνει:

- Τον εντοπισμό των κατάλληλων πηγών δεδομένων, ανάλογα με το άθλημα, τον στόχο της ανάλυσης και τα διαθέσιμα εργαλεία.
- Τη μεθοδολογία καταγραφής, που μπορεί να είναι είτε αυτοματοποιημένη (μέσω συστημάτων) είτε χειροκίνητη (π.χ. από τεχνικό επιτελείο ή αναλυτές).
- Την αξιολόγηση της αξιοπιστίας των δεδομένων, ώστε να διασφαλίζεται η ποιότητά τους και να αποφεύγονται λανθασμένα ή παραπλανητικά συμπεράσματα.

Η σημασία αυτής της φάσης είναι καθοριστική, καθώς κάθε επόμενο βήμα (όπως η επεξεργασία ή η ανάλυση) εξαρτάται από το κατά πόσο τα αρχικά δεδομένα είναι επαρκή, ακριβή και κατάλληλα για χρήση. Τα σφάλματα στη συλλογή (όπως ελλιπείς εγγραφές, λάθη μέτρησης ή ασυνέπεια στις μεθόδους καταγραφής) επηρεάζουν άμεσα την αξιοπιστία της συνολικής ανάλυσης.

Η συλλογή δεδομένων αποτελεί το θεμέλιο για τις διαδικασίες που ακολουθούν:

- Στο 3.2, θα παρουσιαστούν οι μέθοδοι επεξεργασίας και καθαρισμού αυτών των δεδομένων, ώστε να αποκτήσουν τελική μορφή ανάλυσης.
- Στο 3.3, θα εξετάσουμε τα ψηφιακά εργαλεία και τις γλώσσες προγραμματισμού που χρησιμοποιούνται για την αποθήκευση, μετατροπή και αξιοποίηση των δεδομένων.

- Στο 3.4, θα αναλυθούν οι μέθοδοι στατιστικής ανάλυσης και μηχανικής μάθησης, οι οποίες χρησιμοποιούνται για την εξαγωγή συμπερασμάτων και προβλέψεων βάσει των δεδομένων αυτών.

Η ροή της εργασίας, επομένως, ξεκινά με τη συλλογή και συνεχίζει σε μια λογική αλυσίδα βημάτων που καθιστούν δυνατή την αξιόπιστη εφαρμογή της ανάλυσης δεδομένων στον αθλητισμό.

3.2 Επεξεργασία και καθαρισμός δεδομένων

Αφού ολοκληρωθεί η φάση της συλλογής των δεδομένων, όπως περιγράφηκε στο προηγούμενο υποκεφάλαιο, ακολουθεί ένα ιδιαίτερα κρίσιμο βήμα: η επεξεργασία και ο καθαρισμός των δεδομένων. Αυτή η διαδικασία εξασφαλίζει ότι τα δεδομένα που θα χρησιμοποιηθούν για στατιστική ανάλυση ή μηχανική μάθηση είναι αξιόπιστα, συνεκτικά και απαλλαγμένα από σφάλματα. Όπως είναι γνωστό, η ποιότητα της ανάλυσης εξαρτάται άμεσα από την ποιότητα των δεδομένων που τη στηρίζουν· επομένως, η φάση αυτή δεν πρέπει να υποτιμάται.

Κατά την καταγραφή αθλητικών δεδομένων —είτε πρόκειται για στατιστικά παικτών, είτε για μετρήσεις φυσικής κατάστασης— είναι συχνό φαινόμενο να παρουσιάζονται ελλείψεις, σφάλματα εισαγωγής, διπλοεγγραφές ή τιμές εκτός λογικού εύρους. Για παράδειγμα, ένας παίκτης μπορεί να εμφανίζεται με ακραίες τιμές σε ταχύτητα ή καρδιακό ρυθμό, που δεν ανταποκρίνονται στην πραγματικότητα. Επίσης, οι διαφορετικές πηγές δεδομένων ενδέχεται να χρησιμοποιούν διαφορετικές μονάδες μέτρησης ή μορφές καταγραφής, γεγονός που απαιτεί ομογενοποίηση των εγγραφών για να καταστούν συγκρίσιμες.

Η διαδικασία καθαρισμού των δεδομένων περιλαμβάνει εντοπισμό και διαχείριση των ελλιπών τιμών, είτε με απαλοιφή των αντίστοιχων γραμμών, είτε με συμπλήρωσή τους μέσω στατιστικών μεθόδων (όπως ο μέσος όρος ή η παλινδρόμηση). Παράλληλα, γίνεται μετατροπή των μονάδων μέτρησης όπου απαιτείται, όπως για παράδειγμα από μίλια σε χιλιόμετρα ή από λεπτά σε δευτερόλεπτα, ώστε τα δεδομένα να είναι συγκρίσιμα μεταξύ τους. Επιπλέον, αφαιρούνται οι διπλοεγγραφές και τυποποιούνται τα πεδία με κατηγορίες, όπως τα ονόματα ομάδων, οι ημερομηνίες ή οι θέσεις παικτών.

Η επεξεργασία αυτή μπορεί να πραγματοποιηθεί μέσω διαφόρων εργαλείων. Σε απλές περιπτώσεις, μπορεί να γίνει με τη χρήση Excel ή Google Sheets. Ωστόσο, για μεγαλύτερα ή πιο σύνθετα σύνολα δεδομένων, αξιοποιούνται ευρύτερα προγραμματιστικά εργαλεία όπως η γλώσσα Python με τις βιβλιοθήκες pandas και numpy, ή η γλώσσα R, που είναι ιδιαίτερα δημοφιλής στην ακαδημαϊκή κοινότητα για στατιστική ανάλυση. Επιπλέον, η γλώσσα SQL χρησιμοποιείται για την άντληση και φιλτράρισμα δεδομένων από βάσεις δεδομένων. Η επιλογή του κατάλληλου εργαλείου εξαρτάται από την πολυπλοκότητα της ανάλυσης και τον όγκο των δεδομένων.

3.3 Χρήση εργαλείων και οι γλώσσες προγραμματισμού

Όπως αναφέρθηκε και στο προηγούμενο υποκεφάλαιο, η επεξεργασία και ανάλυση των αθλητικών δεδομένων προϋποθέτει τη χρήση κατάλληλων εργαλείων και τεχνολογιών. Με τον ολοένα αυξανόμενο όγκο δεδομένων στον αθλητισμό, καθίσταται απαραίτητη η αξιοποίηση προγραμματιστικών γλωσσών και λογισμικών που προσφέρουν ευελιξία, αυτοματισμό και δυνατότητες βαθύτερης ανάλυσης. Η επιλογή του κατάλληλου εργαλείου εξαρτάται τόσο από τη φύση και την ποσότητα των δεδομένων όσο και από τους στόχους της εκάστοτε ανάλυσης.

Ένα από τα πιο διαδεδομένα εργαλεία είναι το Microsoft Excel, το οποίο παραμένει ιδιαίτερα χρήσιμο για βασικές στατιστικές λειτουργίες, γρήγορη οπτικοποίηση, φιλτράρισμα και ταξινόμηση δεδομένων. Παρά τον περιορισμό του σε μεγάλους όγκους δεδομένων ή πολύπλοκες διαδικασίες, το Excel αποτελεί μια προσβάσιμη λύση για αρχική ανάλυση και δημιουργία απλών γραφημάτων ή συγκεντρωτικών πινάκων. Σε πολλές περιπτώσεις χρησιμοποιείται για προκαταρκτική επεξεργασία πριν τα δεδομένα μεταφερθούν σε πιο εξειδικευμένα περιβάλλοντα.

Για πιο σύνθετες και επιστημονικές αναλύσεις, οι γλώσσες Python και R κατέχουν κυρίαρχη θέση στον χώρο της ανάλυσης δεδομένων. Η Python, λόγω της απλότητας και ευελιξίας της, είναι ιδιαίτερα δημοφιλής. Με βιβλιοθήκες όπως οι pandas για χειρισμό δεδομένων, matplotlib και seaborn για γραφήματα, καθώς και scikit-learn για μοντέλα μηχανικής μάθησης, η Python παρέχει ένα ολοκληρωμένο εργαλείο για την επεξεργασία, την ανάλυση, και την πρόβλεψη βάσει αθλητικών δεδομένων. Επιπλέον, με την προσθήκη βιβλιοθηκών όπως TensorFlow ή PyTorch, είναι δυνατή η ανάπτυξη

προηγμένων μοντέλων τεχνητής νοημοσύνης, όπως νευρωνικά δίκτυα για την πρόβλεψη απόδοσης αθλητών ή τραυματισμών.

Η γλώσσα R, από την άλλη πλευρά, είναι ιδιαίτερα χρήσιμη σε ακαδημαϊκές και ερευνητικές προσεγγίσεις. Ειδικεύεται στη στατιστική ανάλυση και προσφέρει ένα ευρύ φάσμα πακέτων για τη δημιουργία εξειδικευμένων στατιστικών μοντέλων, όπως γραμμική παλινδρόμηση, ανάλυση διασποράς και μοντέλα επιβίωσης. Είναι ιδανική για έργα που απαιτούν λεπτομερή στατιστική ερμηνεία, όπως για παράδειγμα η ανάλυση της φυσικής κατάστασης ενός αθλητή σε βάθος χρόνου ή η σύγκριση αποδόσεων μεταξύ διαφορετικών ομάδων.

Ένα ακόμα σημαντικό εργαλείο είναι η SQL (Structured Query Language), η οποία χρησιμοποιείται κυρίως για τη διαχείριση και εξαγωγή δεδομένων από σχεσιακές βάσεις δεδομένων. Σε αθλητικές οργανώσεις, ακαδημίες ή επαγγελματικές ομάδες, όπου τα δεδομένα αποθηκεύονται σε μεγάλες βάσεις, η SQL καθίσταται απαραίτητη για την αποδοτική αναζήτηση, οργάνωση και συνδυασμό δεδομένων. Μέσα από queries, μπορούν να αντληθούν πολύτιμες πληροφορίες, όπως οι μέσες επιδόσεις αθλητών ανά σεζόν, οι συσχετισμοί μεταξύ τύπων τραυματισμών και προπονητικών φορτίων, ή οι διαφορές απόδοσης ανά αντίπαλο.

Πέρα από τις γλώσσες αυτές, ολοένα και περισσότερες εφαρμογές χρησιμοποιούν BI εργαλεία (Business Intelligence) όπως το Tableau και το Power BI, τα οποία επιτρέπουν τη δημιουργία διαδραστικών dashboards. Μέσω αυτών, οι προπονητές και οι αναλυτές μπορούν να παρακολουθούν τις εξελίξεις σε πραγματικό χρόνο, να συγκρίνουν παίκτες ή ομάδες και να λαμβάνουν αποφάσεις βασισμένες σε απτά δεδομένα και οπτικές αναπαραστάσεις.

Η κατανόηση και σωστή χρήση αυτών των εργαλείων και γλωσσών προγραμματισμού είναι κρίσιμη για την επιτυχία οποιασδήποτε μορφής ανάλυσης στον αθλητισμό. Η επιλογή τους δεν εξαρτάται μόνο από τις τεχνικές δυνατότητες του κάθε μέσου, αλλά και από το είδος των δεδομένων, τον χρόνο που υπάρχει για την ανάλυση και το επίπεδο τεχνογνωσίας των χρηστών. Σε κάθε περίπτωση, η τεχνολογική κατάρτιση αναδεικνύεται ως ένα βασικό προσόν για κάθε επαγγελματία ή ερευνητή που επιθυμεί να χρησιμοποιήσει την ανάλυση δεδομένων ως μέσο βελτιστοποίησης της απόδοσης και της στρατηγικής στον αθλητισμό.

3.4 Μέθοδοι στατιστικής ανάλυσης και μηχανικής μάθησης στον αθλητισμό

Η μηχανική μάθηση (machine learning) αποτελεί έναν δυναμικά αναπτυσσόμενο υποτομέα της τεχνητής νοημοσύνης και έχει ως βασικό στόχο την κατανόηση της δομής των δεδομένων και τη σύνδεσή τους με μοντέλα που μπορούν να ερμηνευθούν και να αξιοποιηθούν. Ενώ η τεχνητή νοημοσύνη στοχεύει στην προσομοίωση ανθρώπινης σκέψης και λήψης αποφάσεων, η μηχανική μάθηση επικεντρώνεται στην ικανότητα των υπολογιστικών συστημάτων να "μαθαίνουν" μέσα από τα δεδομένα και να βελτιώνουν την απόδοσή τους χωρίς να έχουν προγραμματιστεί ρητά για κάθε εργασία. Ο όρος "μηχανική μάθηση" επινοήθηκε από τον Arthur Samuel το 1959, και αργότερα αποσαφηνίστηκε μαθηματικά από τον Tom Mitchell το 1997, ο οποίος όρισε πως ένα υπολογιστικό σύστημα μαθαίνει όταν η απόδοσή του σε μία εργασία T μετρήσιμη από ένα μέτρο P βελτιώνεται με την εμπειρία E. Η ουσία της μηχανικής μάθησης βρίσκεται στη δυνατότητα αξιοποίησης τεράστιων όγκων δεδομένων, στην αυξανόμενη υπολογιστική ισχύ και στη συνεχή βελτίωση των αλγορίθμων.

Η διαδικασία της μηχανικής μάθησης περιλαμβάνει διακριτά και καθορισμένα βήματα, εκ των οποίων τα πρώτα είναι καθοριστικής σημασίας. Στο πρώτο βήμα το θεμέλιο κάθε επιτυχημένου μοντέλου είναι η ύπαρξη κατάλληλων και ποιοτικών δεδομένων. Η επιλογή των δεδομένων που θα χρησιμοποιηθούν στην εκπαίδευση του μοντέλου επηρεάζει άμεσα την αποτελεσματικότητά του. Εσφαλμένη επιλογή χαρακτηριστικών, περιορισμένο εύρος καταγραφών ή λανθασμένη κατηγοριοποίηση μπορούν να οδηγήσουν σε ανεπαρκή ή προκατειλημμένα αποτελέσματα. Για παράδειγμα, σε ένα μοντέλο που καλείται να αναγνωρίσει αθλητικές δραστηριότητες με βάση την ταχύτητα, τα δεδομένα που προέρχονται από ένα περιορισμένο δείγμα πληθυσμού μπορεί να οδηγήσουν σε εσφαλμένες εκτιμήσεις (Constantinou & Fenton, 2013).

Ακολουθεί επόμενο βήμα: Προετοιμασία των δεδομένων, κατά το οποίο τα δεδομένα καθαρίζονται και οργανώνονται κατάλληλα για να είναι έτοιμα προς εκπαίδευση. Σε αυτό το στάδιο, αφαιρούνται τα διπλότυπα, διορθώνονται λάθη, αντιμετωπίζονται οι ελλιπείς τιμές και πραγματοποιούνται μετατροπές στους τύπους των δεδομένων. Η τυχαιοποίηση της σειράς των καταγραφών είναι σημαντική ώστε να αποφευχθεί η

επιρροή της διάταξης των δεδομένων στο αποτέλεσμα της εκπαίδευσης. Εξίσου κρίσιμη είναι η διάσπαση των δεδομένων σε δύο σύνολα: ένα για την εκπαίδευση του μοντέλου (συνήθως 80%) και ένα για την αξιολόγησή του (20%). Αυτή η διάκριση διασφαλίζει ότι η απόδοση του μοντέλου θα μετρηθεί σε νέα, "άγνωστα" δεδομένα, προσομοιώνοντας με μεγαλύτερη ακρίβεια τις πραγματικές συνθήκες εφαρμογής.

Το βήμα 3: Επιλογή μοντέλου επικεντρώνεται στην εύρεση του κατάλληλου τύπου αλγορίθμου για το πρόβλημα που μελετάται. Ανάλογα με τη φύση των δεδομένων και το ζητούμενο αποτέλεσμα (π.χ. ταξινόμηση, παλινδρόμηση, σύσταση), επιλέγεται ένα μοντέλο που να ανταποκρίνεται καλύτερα στις ανάγκες του έργου. Υπάρχουν αλγόριθμοι που είναι πιο αποτελεσματικοί στην επεξεργασία κειμένου (όπως οι Naive Bayes ή τα RNNs), ενώ άλλοι είναι καταλληλότεροι για εικόνες (όπως τα CNNs). Η σωστή επιλογή μοντέλου απαιτεί γνώση τόσο του αλγορίθμου όσο και του πεδίου εφαρμογής, καθώς και πειραματισμό για τον εντοπισμό της βέλτιστης λύσης.

Σε αντίθεση με τον παραδοσιακό προγραμματισμό, όπου οι κανόνες καθορίζονται εξ αρχής από τον προγραμματιστή, στη μηχανική μάθηση οι κανόνες αυτοί "ανακαλύπτονται" από το σύστημα, μέσω της επεξεργασίας των δεδομένων. Αυτή η ιδιαιτερότητα της μηχανικής μάθησης την καθιστά ιδιαίτερα πολύτιμη για εφαρμογές όπου οι κανόνες είναι περίπλοκοι, δυσδιάκριτοι ή μεταβαλλόμενοι, όπως στην ανάλυση αθλητικής απόδοσης, την πρόβλεψη τραυματισμών ή τη δημιουργία εξατομικευμένων προγραμμάτων προπόνησης. Έτσι, η σύνδεση μεταξύ της έννοιας της μηχανικής μάθησης και των πρώτων της σταδίων (συλλογή, καθαρισμός και επιλογή μοντέλου) αποδεικνύεται κρίσιμη για την επιτυχή εφαρμογή της στον χώρο του αθλητισμού και όχι μόνο.

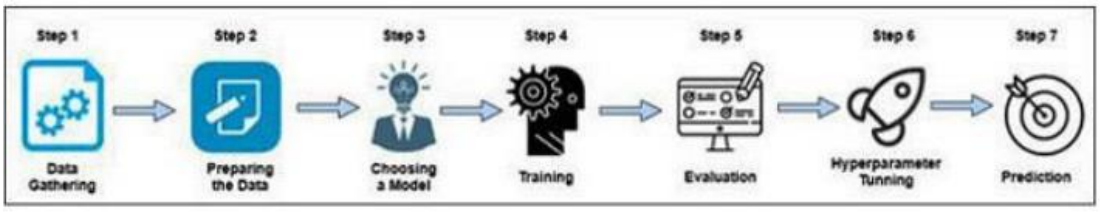
Η εκπαίδευση του μοντέλου αποτελεί το κεντρικό στάδιο της μηχανικής μάθησης. Σε αυτή τη φάση, το μοντέλο «μαθαίνει» να αναγνωρίζει πρότυπα και συσχετίσεις μέσα στα δεδομένα εκπαίδευσης. Με την αρχική τυχαία τοποθέτηση παραμέτρων, το μοντέλο παράγει προβλέψεις που συγκρίνονται με τις πραγματικές τιμές-στόχους. Με τη βοήθεια μαθηματικών αλγορίθμων, οι παράμετροι προσαρμόζονται ώστε να μειωθεί το σφάλμα μεταξύ πρόβλεψης και πραγματικότητας. Η διαδικασία αυτή επαναλαμβάνεται πολλές φορές (βήματα εκπαίδευσης ή epochs), βελτιώνοντας

σταδιακά την απόδοση του μοντέλου. Ο αλγόριθμος «μαθαίνει» όπως ένα παιδί που μαθαίνει να οδηγεί ποδήλατο: αρχικά κάνει λάθη, αλλά σταδιακά βελτιώνεται.

Μετά την εκπαίδευση, το μοντέλο πρέπει να αξιολογηθεί σε νέα, άγνωστα δεδομένα, για να διαπιστωθεί εάν όντως έχει μάθει να γενικεύει και να λειτουργεί σωστά σε πραγματικές συνθήκες. Για το σκοπό αυτό χρησιμοποιείται το σύνολο δεδομένων αξιολόγησης (test set), που δεν έχει χρησιμοποιηθεί κατά την εκπαίδευση. Μετρώνται διάφορα κριτήρια απόδοσης (όπως ακρίβεια, ανάκληση, F1-score) ανάλογα με το πρόβλημα. Η σωστή αξιολόγηση είναι κρίσιμη, ειδικά όταν το μοντέλο πρόκειται να χρησιμοποιηθεί σε εμπορικές εφαρμογές ή να λάβει αποφάσεις με σημαντικές επιπτώσεις, καθώς μια υπερεκπαιδευμένη ή υποαποδοτική λύση μπορεί να οδηγήσει σε λανθασμένες προβλέψεις και απώλεια εμπιστοσύνης.

Οι υπερπαράμετροι είναι ρυθμίσεις που επηρεάζουν την εκπαίδευση και τη δομή του μοντέλου, όπως ο ρυθμός μάθησης, το μέγεθος του batch, ο αριθμός των επιπέδων ενός νευρωνικού δικτύου κ.ά. Αυτές οι παράμετροι δεν μαθαίνονται αυτόματα, αλλά πρέπει να οριστούν πριν ή κατά τη διάρκεια της εκπαίδευσης. Ο συντονισμός υπερπαραμέτρων στοχεύει στην εύρεση του βέλτιστου συνδυασμού τους ώστε να μεγιστοποιηθεί η απόδοση του μοντέλου. Αυτό γίνεται μέσω διαδικασιών όπως η αναζήτηση πλέγματος (grid search) ή τυχαία αναζήτηση (random search), ενώ πιο προηγμένες τεχνικές χρησιμοποιούν αλγορίθμους βελτιστοποίησης. Ο συντονισμός είναι απαραίτητος, ειδικά σε ανταγωνιστικά περιβάλλοντα, καθώς το πιο ακριβές και αποδοτικό μοντέλο υπερισχύει (Hvattum & Arntzen, 2010).

Το τελικό στάδιο της μηχανικής μάθησης είναι η πρόβλεψη, όπου το εκπαιδευμένο και βελτιστοποιημένο μοντέλο χρησιμοποιείται για να παράγει αποτελέσματα σε νέα δεδομένα που δεν έχει ξαναδεί. Σε αυτό το στάδιο αποδεικνύεται η πραγματική αξία της μηχανικής μάθησης, καθώς το μοντέλο εφαρμόζεται σε πραγματικές εφαρμογές, όπως η αναγνώριση εικόνων, η πρόβλεψη συμπεριφορών ή η ταξινόμηση κειμένων. Η ποιότητα των προβλέψεων εξαρτάται άμεσα από την ποιότητα των δεδομένων, την ακρίβεια του μοντέλου και το σωστό συντονισμό όλων των παραμέτρων στη διάρκεια της διαδικασίας.



Εικόνα 12: Βήματα μηχανικής μάθησης

Με την ολοκλήρωση αυτών των βημάτων, το σύστημα μηχανικής μάθησης είναι έτοιμο να λειτουργήσει ως αυτόνομο εργαλείο ανάλυσης και λήψης αποφάσεων, προσφέροντας σημαντική προστιθέμενη αξία σε διάφορους τομείς, όπως ο αθλητισμός, η υγεία, η οικονομία και πολλά ακόμη.

Τα βήματα συνοψίζοντας, με σειρά, είναι: Συλλογή δεδομένων, επεξεργασία και προετοιμασία των δεδομένων, επιλογή μοντέλου, εκπαίδευση του machine learning, αξιολόγηση, συντονισμός υπερπαραμέτρων και τελικώς, πρόβλεψη (FanSided, 2017).

Πλεονεκτήματα της Μηχανικής Μάθησης

1. Αναγνώριση τάσεων και μοτίβων: Μπορεί να επεξεργαστεί μεγάλες ποσότητες δεδομένων και να εντοπίσει κρυφές δομές και μοτίβα που είναι δύσκολο να εντοπίσει ένας άνθρωπος. Για παράδειγμα, ένας ιστότοπος ηλεκτρονικού εμπορίου μπορεί να κατανοήσει τη συμπεριφορά των χρηστών και να προτείνει σχετικά προϊόντα ή διαφημίσεις.

2. Αυτοματοποίηση χωρίς ανθρώπινη παρέμβαση: Επιτρέπει την αυτοματοποίηση σύνθετων εργασιών λήψης αποφάσεων, μειώνοντας την ανάγκη για συνεχή ανθρώπινη παρέμβαση, κάτι που αυξάνει την αξιοπιστία και την αποδοτικότητα.

3. Επεξεργασία οποιουδήποτε τύπου δεδομένων: Μπορεί να χειριστεί διαφορετικούς τύπους και μεγέθη δεδομένων που δύσκολα διαχειρίζονται παραδοσιακά συστήματα, καθιστώντας την πολύ ευέλικτη.

4. Συνεχής βελτίωση με την εμπειρία: Καθώς εκπαιδεύεται με νέα δεδομένα, το μοντέλο βελτιώνει σταδιακά την ακρίβεια και την απόδοσή του, μαθαίνοντας από τα λάθη και τις επιτυχίες του, π.χ. στην πρόβλεψη καιρού.

Μειονεκτήματα της Μηχανικής Μάθησης

1. Υψηλή πιθανότητα σφάλματος: Λόγω της φύσης των μεγάλων δεδομένων, είναι πιθανό να υπάρχουν σφάλματα ή ανεπιθύμητα μοτίβα που δυσκολεύουν την ακριβή εκπαίδευση. Η διόρθωσή τους μπορεί να είναι χρονοβόρα ή και αδύνατη σε ορισμένες περιπτώσεις.

2. Ανάγκη για μεγάλα και ποιοτικά σύνολα δεδομένων: Για να εκπαιδευτεί σωστά ένα μοντέλο, απαιτούνται μεγάλες ποσότητες δεδομένων υψηλής ποιότητας. Τα δεδομένα συχνά πρέπει να ενημερώνονται συνεχώς, κάτι που μπορεί να περιπλέξει τη διαδικασία.

3. Μεγάλος χρόνος εκπαίδευσης και ανάπτυξης: Η εκπαίδευση των μοντέλων μηχανικής μάθησης χρειάζεται αρκετό χρόνο ώστε να επιτευχθεί ο επιθυμητός βαθμός ακρίβειας και συνέπειας.

4. Υψηλές υπολογιστικές απαιτήσεις: Η διαδικασία συχνά απαιτεί ισχυρούς υπολογιστικούς πόρους, όπως GPU και μεγάλης χωρητικότητας μνήμη, κάτι που μπορεί να αυξήσει το κόστος και την πολυπλοκότητα.

3.5 Εφαρμογές μηχανικής μάθησης στον αθλητισμό

Παράδειγμα: Πρόβλεψη του σκορ ενός αγώνα

Η πρόβλεψη σκορ σε αθλητικούς αγώνες μέσω μηχανικής μάθησης είναι μια πολύ ενδιαφέρουσα και απαιτητική διαδικασία που βασίζεται στην ανάλυση μεγάλου όγκου δεδομένων και στην εκπαίδευση αλγορίθμων για να μάθουν να προβλέπουν το αποτέλεσμα με βάση ιστορικές και τρέχουσες πληροφορίες (Tavares, 2018).

Βήματα για την πρόβλεψη σκορ με μηχανική μάθηση

1. Συλλογή δεδομένων: Αρχικά, συλλέγονται μεγάλα και ποικίλα δεδομένα που σχετίζονται με το άθλημα και τους αγώνες. Αυτά μπορεί να περιλαμβάνουν:
    - Ιστορικά σκορ αγώνων
    - Στατιστικά αθλητών και ομάδων (π.χ. μέσος όρος γκολ, ασίστ, ποσοστό επιτυχίας)
    - Στοιχεία φυσικής κατάστασης και τραυματισμών

- Παράγοντες περιβάλλοντος (καιρικές συνθήκες, γήπεδο, ώρα αγώνα)
- Στοιχεία τακτικής και αλλαγές στη σύνθεση της ομάδας

2. Επεξεργασία και προετοιμασία δεδομένων: Τα δεδομένα καθαρίζονται, τυποποιούνται και προετοιμάζονται για να γίνουν κατανοητά από τους αλγόριθμους. Μπορεί να χρειαστεί να αντιμετωπιστούν ελλιπή ή ασαφή δεδομένα, να γίνουν κανονικοποιήσεις ή να δημιουργηθούν νέα χαρακτηριστικά (feature engineering).
3. Επιλογή και εκπαίδευση μοντέλου: Επιλέγονται κατάλληλοι αλγόριθμοι μηχανικής μάθησης, όπως:
   - Αλγόριθμοι ταξινόμησης ή παλινδρόμησης(π.χ. Random Forest, Gradient Boosting, Support Vector Machines, Νευρωνικά Δίκτυα)
   - Μοντέλα ακολουθίας ή χρονικής σειράς (π.χ. LSTM, RNN) που μπορούν να προβλέψουν το σκορ λαμβάνοντας υπόψη τη χρονική εξέλιξη των δεδομένων. Το μοντέλο εκπαιδεύεται με τα ιστορικά δεδομένα για να μάθει τη σχέση μεταξύ των χαρακτηριστικών και του τελικού σκορ.
4. Αξιολόγηση μοντέλου: Το μοντέλο δοκιμάζεται σε νέα δεδομένα (που δεν έχουν χρησιμοποιηθεί στην εκπαίδευση) για να εκτιμηθεί η ακρίβεια και η αξιοπιστία της πρόβλεψης. Χρησιμοποιούνται μετρικές όπως το Mean Squared Error (MSE), accuracy, ή F1-score, ανάλογα με το είδος του προβλήματος (πρόβλεψη ακριβούς σκορ ή νικητή).
5. Πρόβλεψη: Με βάση τα τρέχοντα δεδομένα (προηγούμενοι αγώνες, κατάσταση ομάδων, τραυματισμοί κ.ά.), το μοντέλο προβλέπει το πιθανό σκορ για τον επερχόμενο αγώνα.

Προκλήσεις και παράγοντες που επηρεάζουν την πρόβλεψη

- Πολυπλοκότητα αθλήματος: Η δυναμική των αθλημάτων είναι σύνθετη, με πολλούς παράγοντες που επηρεάζουν το αποτέλεσμα (π.χ. τύχη, ψυχολογία).

- Ποιότητα και ποσότητα δεδομένων: Χρειάζονται μεγάλα και αξιόπιστα δεδομένα για να είναι το μοντέλο αποτελεσματικό.
- Απρόβλεπτα γεγονότα: Τραυματισμοί, αλλαγές προπονητών ή καιρικές συνθήκες μπορούν να αλλοιώσουν την πρόβλεψη.

Με σωστή χρήση και συνδυασμό δεδομένων και αλγορίθμων, η μηχανική μάθηση μπορεί να δώσει πολύτιμες ενδείξεις και προβλέψεις που υποστηρίζουν προπονητές, αναλυτές και στοιχηματικές εταιρείες στον αθλητικό χώρο.

Η μηχανική μάθηση, ως ένας από τους πιο δυναμικούς τομείς της τεχνητής νοημοσύνης, έχει βρει ευρεία εφαρμογή στον αθλητισμό, προσφέροντας εργαλεία για την ανάλυση και βελτιστοποίηση της απόδοσης των αθλητών. Μέσω αλγορίθμων που «μαθαίνουν» από τα δεδομένα, η μηχανική μάθηση μπορεί να εντοπίσει πρότυπα και τάσεις που δεν είναι άμεσα ορατά στον ανθρώπινο παρατηρητή, επιτρέποντας πιο τεκμηριωμένες αποφάσεις κατά την προπόνηση και τους αγώνες.

Ένα σημαντικό πεδίο εφαρμογής της μηχανικής μάθησης είναι η πρόβλεψη τραυματισμών. Μέσω της ανάλυσης ιστορικών δεδομένων, όπως φυσική κατάσταση, φορτίο προπόνησης και προηγούμενοι τραυματισμοί, τα μοντέλα μηχανικής μάθησης μπορούν να αξιολογήσουν τον κίνδυνο τραυματισμού και να ειδοποιήσουν προπονητές και αθλητές εγκαίρως. Αυτό επιτρέπει την έγκαιρη προσαρμογή των προγραμμάτων προπόνησης και την πρόληψη σοβαρών προβλημάτων υγείας.

Επιπλέον, η μηχανική μάθηση συμβάλλει στη βελτίωση της στρατηγικής και τακτικής στα αθλήματα. Αναλύοντας δεδομένα αγώνων και απόδοσης, οι αλγόριθμοι μπορούν να υποδείξουν βέλτιστες κινήσεις, αλλαγές σχημάτων και τακτικές προσαρμογές σε πραγματικό χρόνο. Αυτή η δυνατότητα έχει εφαρμογές σε ομαδικά αθλήματα όπως το ποδόσφαιρο, το μπάσκετ και το χάντμπολ, όπου η στρατηγική παίζει καθοριστικό ρόλο στην έκβαση του αγώνα.

Η μηχανική μάθηση χρησιμοποιείται επίσης για την εξατομίκευση των προγραμμάτων προπόνησης. Με τη συλλογή δεδομένων από αισθητήρες, ρολόγια και άλλες φορητές συσκευές, τα μοντέλα αναλύουν τη φυσική κατάσταση και την απόδοση κάθε αθλητή και δημιουργούν προσαρμοσμένα προγράμματα που μεγιστοποιούν τα αποτελέσματα ενώ μειώνουν τον κίνδυνο υπερπροπόνησης (Igiri, 2015).

Τέλος, η μηχανική μάθηση παίζει ρόλο στην ανάλυση βίντεο και την αξιολόγηση τεχνικής. Μέσω της αναγνώρισης προτύπων και της επεξεργασίας εικόνας, τα συστήματα μπορούν να εντοπίσουν λάθη στην τεχνική, να αναλύσουν κινήσεις και να προσφέρουν ανατροφοδότηση στους αθλητές και τους προπονητές για βελτιώσεις, συμβάλλοντας σημαντικά στην τεχνική εκμάθηση και ανάπτυξη (Tilp & Schrapf, 2015).

3.6 Η είσοδος και η εφαρμογή της τεχνητής νοημοσύνης στον αθλητισμό

Η τεχνητή νοημοσύνη (Artificial Intelligence – AI) έχει μπει δυναμικά στον χώρο του αθλητισμού, μεταμορφώνοντας όχι μόνο την ίδια την αγωνιστική διαδικασία αλλά και το ευρύτερο οικοσύστημα της αθλητικής βιομηχανίας. Η αξιοποίηση της AI οδηγεί σε πιο επαγγελματική, επιστημονική και εξατομικευμένη προσέγγιση σε πολλαπλά επίπεδα.

Βελτίωση της αγωνιστικής απόδοσης και υποστήριξη προπονητών

Η εφαρμογή της AI επιτρέπει την ακριβή παρακολούθηση της φυσικής κατάστασης των αθλητών σε πραγματικό χρόνο — πριν, κατά τη διάρκεια και μετά τους αγώνες. Μέσω της συνεχούς συλλογής και ανάλυσης δεδομένων (μεγάλα δεδομένα), οι προπονητές λαμβάνουν υποστήριξη για να προσαρμόζουν τις τακτικές τους άμεσα, βελτιώνοντας τις στρατηγικές και αυξάνοντας την ανταγωνιστικότητα των αθλητών. Παράλληλα, η AI επιτρέπει τη δημιουργία εξατομικευμένων μοντέλων προπόνησης που λαμβάνουν υπόψη τις ατομικές ανάγκες και ιδιαιτερότητες κάθε αθλητή.

Αναβάθμιση της αθλητικής βιομηχανίας

Ο αθλητικός τομέας, ως βιομηχανία με σημαντική προοπτική ανάπτυξης, ωφελείται από την εφαρμογή των τεχνολογιών AI στην οργάνωση, τη χρηματοδότηση και την καινοτομία. Η τεχνητή νοημοσύνη συμβάλλει στη βελτιστοποίηση της χρήσης κεφαλαίων και στη δημιουργία νέων επιχειρηματικών μοντέλων που βασίζονται σε επιστημονικά δεδομένα και έξυπνες τεχνολογίες. Αυτό βοηθά στη μετάβαση της παραδοσιακής αθλητικής βιομηχανίας προς μια πιο «έξυπνη» και ανταγωνιστική μορφή, ενισχύοντας τη θέση της στον παγκόσμιο χάρτη (Vanschoren et al., 2012).

Προώθηση της εξατομικευμένης φυσικής αγωγής

Η AI, μαζί με άλλες τεχνολογίες πληροφορικής και μεγάλα δεδομένα, προωθεί την εξατομικευμένη φυσική αγωγή σε σχολεία και εκπαιδευτικά ιδρύματα. Δημιουργείται ένα νέο επιστημονικό οικοσύστημα που επιτρέπει στους μαθητές να συμμετέχουν σε στοχευμένες αθλητικές δραστηριότητες και κατάρτιση, ενώ οι εκπαιδευτικοί λαμβάνουν πιο ακριβή και επιστημονική καθοδήγηση. Επιπλέον, διευκολύνεται η αρμονική συνεργασία μεταξύ σχολείων, κοινωνίας και οικογένειας μέσω της ανταλλαγής δεδομένων.

Στρατηγικές για την εφαρμογή της AI στον αθλητισμό

Η επιτυχής εφαρμογή της AI στον αθλητισμό απαιτεί συντονισμένες πολιτικές και οικονομική υποστήριξη από τις κυβερνήσεις και τα αθλητικά τμήματα. Είναι απαραίτητο να αποσαφηνιστούν οι αρμοδιότητες των εμπλεκόμενων φορέων και να δημιουργηθούν κατάλληλα κίνητρα και πλαίσια για την ενθάρρυνση της έρευνας και της καινοτομίας. Εθνικά σχέδια και πολιτικές οφείλουν να καθοδηγούν την ανάπτυξη της αθλητικής AI, ενώ πειραματικές βάσεις και ερευνητικά κέντρα πρέπει να υποστηρίζονται οικονομικά και θεσμικά.

Ηθικά ζητήματα και πρότυπα εποπτείας

Η ανάπτυξη και χρήση της τεχνητής νοημοσύνης στον αθλητισμό εγείρει σημαντικά ηθικά και ρυθμιστικά ζητήματα. Είναι απαραίτητη η διαμόρφωση εθνικών προτύπων για την εφαρμογή της AI, με σαφείς κανόνες που διασφαλίζουν την αθλητική δικαιοσύνη και την προστασία των δικαιωμάτων των χρηστών. Η δημιουργία δικτύων εποπτείας και η συνεργασία μεταξύ κρατικών οργάνων και αθλητικών φορέων είναι κρίσιμη για την ομαλή και δίκαιη λειτουργία των ευφυών αθλητικών συστημάτων. Τεχνολογίες όπως το blockchain μπορούν να ενισχύσουν την ασφάλεια και την ιδιωτικότητα στα δεδομένα, προωθώντας τη θετική ανάπτυξη του έξυπνου αθλητισμού.

ΚΕΦΑΛΑΙΟ 4: CASE STUDIES

4.1 Brentford Football Club

Η Brentford Football Club αποτελεί ένα από τα πιο χαρακτηριστικά παραδείγματα αξιοποίησης ανάλυσης δεδομένων στο σύγχρονο ποδόσφαιρο, αναδεικνύοντας πώς μια μικρομεσαία ομάδα μπορεί να σταθεί ανταγωνιστικά απέναντι σε αντιπάλους με πολλαπλάσιους πόρους. Η φιλοσοφία της ομάδας βασίζεται στη στρατηγική του «Moneyball», προσαρμοσμένη στο ευρωπαϊκό ποδοσφαιρικό πλαίσιο, με στόχο την ανατροπή των ανισοτήτων μέσω καινοτομίας και ορθολογικής χρήσης της πληροφορίας (Reddit, n.d.).



Εικόνα 13: Ο ποδοσφαιριστής της Brentford Ivan Toney, από τις ακαδημίες της

Η μετάβαση της Brentford από τις χαμηλές κατηγορίες στην Premier League δεν αποτελεί τυχαίο γεγονός, αλλά προϊόν μακροχρόνιου σχεδιασμού, υπό την καθοδήγηση του ιδιοκτήτη Matthew Benham. Ο Benham, επαγγελματίας στον χώρο του στοιχηματισμού και ιδρυτής της εταιρείας Smartodds, εισήγαγε μια νέα κουλτούρα λήψης αποφάσεων που βασίζεται σε στατιστική ανάλυση, μαθηματικά μοντέλα και προγνωστικούς αλγόριθμους. Στελέχη με εμπειρία στον χώρο των

δεδομένων όπως οι Rasmus Ankersen και Phil Giles (με διδακτορικό στη στατιστική) ανέλαβαν καίριους ρόλους στην ποδοσφαιρική διεύθυνση του συλλόγου.

Η ανάλυση δεδομένων αγώνων (match analysis) αποτελεί θεμέλιο της αγωνιστικής προσέγγισης της Brentford. Το τεχνικό επιτελείο δεν επικεντρώνεται αποκλειστικά στο αποτέλεσμα, αλλά σε προγνωστικούς δείκτες απόδοσης όπως τα expected goals (xG), οι αποδόσεις κατοχής, οι μεταβάσεις και οι στατικές φάσεις. Η ερμηνεία της απόδοσης βασίζεται στη θεωρία ότι το αποτέλεσμα ενός ποδοσφαιρικού αγώνα μπορεί να επηρεάζεται υπερβολικά από τυχαία γεγονότα, λόγω του χαμηλού αριθμού γκολ. Έτσι, η ομάδα προτιμά να μελετά τις «υποκείμενες μεταβλητές» (underlying metrics), προκειμένου να προβλέψει τη μακροπρόθεσμη πορεία της και να εντοπίσει πραγματικά ισχυρά ή αδύναμα σημεία στην τακτική της. Ο Ankersen (n.d.) υποστηρίζει ότι η επιτυχία συλλόγων όπως η Brentford βασίζεται σε ορθολογική αξιοποίηση δεδομένων.



Εικόνα 14: Brentford FC: Η ομάδα πρότυπο της Premier League

Η αξιολόγηση της απόδοσης των παικτών ενισχύεται μέσω της ενσωμάτωσης δεδομένων φυσικής κατάστασης, τεχνικών χαρακτηριστικών και συμπεριφοράς στο γήπεδο. Η Brentford χρησιμοποιεί συστήματα παρακολούθησης (GPS trackers, video analysis, sensors) για τη συλλογή πληθώρας δεδομένων όπως αποστάσεις, sprint, θερμικούς χάρτες, ενώ παράλληλα αξιολογεί γνωστικές και ψυχολογικές

παραμέτρους. Οι παίκτες αξιολογούνται όχι μόνο με βάση τα παραδοσιακά στατιστικά, αλλά και βάσει της εξέλιξης τους σε προβλεπόμενες τιμές (projected development) μέσα από τα μοντέλα του συλλόγου.

Όσον αφορά τις προβλέψεις αποτελεσμάτων και τη στρατηγική scouting, η Brentford έχει αναπτύξει μοντέλα σύγκρισης ποιότητας πρωταθλημάτων ανά την Ευρώπη και πέραν αυτής, επιτρέποντάς της να εντοπίζει «κρυμμένα διαμάντια» σε αγορές με χαμηλό ανταγωνισμό. Οι μεταγραφές των Maupay, Benrahma και Watkins αποτελούν χαρακτηριστικά παραδείγματα επιτυχημένων προβλέψεων αξίας και δυναμικής, που μετουσιώθηκαν σε αγωνιστική και οικονομική επιτυχία. Ο Ankersen χαρακτηριστικά σημειώνει: «Δεν σου λέει το data ποιον να πάρεις, σου λέει πού να κοιτάξεις».

Η Brentford δεν δίστασε να κλείσει την ακαδημία της το 2016, καθώς θεώρησε ότι το κόστος ήταν δυσανάλογο σε σχέση με την απόδοση. Αντ' αυτού, δημιούργησε τη B-Team, μια ενδιάμεση αναπτυξιακή ομάδα στην οποία συμμετέχουν νεαροί παίκτες που δεν βρήκαν χώρο σε μεγάλες ομάδες. Μέσω αυτής της δομής, το scouting γίνεται πιο στοχευμένο και εξατομικευμένο, ενισχυμένο από δεδομένα.

Η πρόληψη τραυματισμών είναι ακόμη ένας τομέας όπου η Brentford αξιοποιεί την τεχνολογία. Μέσω της συνεχούς παρακολούθησης της φυσιολογικής κατάστασης των παικτών, εντοπίζονται σημάδια κόπωσης ή υπερφόρτωσης που ενδέχεται να οδηγήσουν σε τραυματισμό. Οι προπονήσεις προσαρμόζονται βάσει αυτών των δεδομένων, ενώ η διαχείριση των παικτών γίνεται με μακροπρόθεσμο ορίζοντα, ενισχύοντας τη διαθεσιμότητα του ρόστερ.

Η στρατηγική επιμονή στη λογική και τα δεδομένα φάνηκε και στην απόφαση της διοίκησης να διατηρήσει τον Thomas Frank παρά τις αρχικές του αποτυχίες. Η ομάδα είχε ενδείξεις από τα στατιστικά μοντέλα ότι οι επιδόσεις του δεν αντανακλώνται δίκαια στα αποτελέσματα, κι έτσι αποφάσισε να στηρίξει τον προπονητή. Η μετέπειτα άνοδος στην Premier League επιβεβαίωσε την επιλογή αυτή.

Τέλος, η επιρροή της Brentford δεν περιορίστηκε στο εσωτερικό της. Η ίδια φιλοσοφία εφαρμόστηκε στην FC Midtjylland, τη δανέζικη ομάδα που επίσης ανήκει στον Benham, ενώ άλλες ομάδες όπως η Barnsley έχουν υιοθετήσει ανάλογες μεθοδολογίες. Η Brentford, χωρίς να διαθέτει το οικονομικό μέγεθος ομάδων όπως η

Liverpool, κατάφερε να πρωτοπορήσει στη μετάβαση του ποδοσφαίρου στη νέα εποχή των δεδομένων (Πηγή: https://analyisport.com/insights/what-can-data-do-for-a-football-club/)

Η Brentford αποδεικνύει έμπρακτα πώς ένα data-driven μοντέλο μπορεί να οδηγήσει:

- Μεταγραφική αποτελεσματικότητα με μεγάλες αποκλίσεις ανάμεσα στα κόστη και τα έσοδα.
- Διατήρηση ανταγωνισμού στην Premier League, με σταθερές μεσαίες θέσεις, παρά μικρότερο μπάτζετ.
- Οικονομική βιωσιμότητα, με κερδοφορία ακόμα και σε περιόδους επενδύσεων.
- Ανάδειξη αστεριών με προστιθέμενη αξία, αξιοποιώντας advanced metrics.
- Βασισμένη στη λογική αντί της τύχης απόδοση και επιλογές, ακόμη και παρά το χαμηλό σκορ στο ποδόσφαιρο.



Εικόνα 15: Δεδομένα ποδοσφαιριστή της Brentford – Εφαρμογή του μοντέλου της

4.2 Ομάδες του NBA και ανάλυση δεδομένων

Το NBA αποτελεί μια από τις κορυφαίες επαγγελματικές λίγκες μπάσκετ στον κόσμο, με 30 ομάδες – 29 στις ΗΠΑ και μία στον Καναδά. Η επιτυχία του δεν οφείλεται μόνο στο αθλητικό ταλέντο των παικτών, αλλά και στην προοδευτική αξιοποίηση της

τεχνολογίας και ειδικότερα των big data (Cai, Zhao & Wang, 2021). Η ανάλυση δεδομένων (data analytics) έχει εξελιχθεί σε κρίσιμο εργαλείο για τις ομάδες, τους παίκτες, τους προπονητές και ακόμα και για τις εμπορικές συνεργασίες του NBA. Η υιοθέτηση τέτοιων τεχνολογιών, με πρωτοστάτη τους Golden State Warriors, προσέφερε απτά αποτελέσματα και ενίσχυσε τον ρόλο των δεδομένων στον αθλητισμό.

Η Χρήση της Τεχνολογίας SportVU

Από τη δεκαετία του 1980, το NBA αξιοποιεί τεχνολογίες καταγραφής στατιστικών. Η επανάσταση όμως ήρθε το 2013 με το σύστημα SportVU, το οποίο καταγράφει κινήσεις παικτών και μπάλας 25 φορές το δευτερόλεπτο, χρησιμοποιώντας κάμερες υψηλής ανάλυσης και τεχνολογία οπτικής παρακολούθησης. Μέσα από αυτό, δημιουργούνται αναλυτικά δεδομένα για ενέργειες όπως σουτ, πάσες, κάλυψη χώρου και απόσταση που διανύεται, προσφέροντας μια νέα βάση ανάλυσης για προπονητές και αναλυτές. Η τεχνολογία SportVU επιτρέπει την οπτική παρακολούθηση κινήσεων παικτών (Stats Perform, n.d.).

Εφαρμογές και Επιρροή στις Ομάδες

Η τεχνολογία αυτή δεν άργησε να διαδοθεί. Οι Mavericks ήταν οι πρώτοι που την εφάρμοσαν και έφτασαν στους τελικούς τη σεζόν 2005-06. Μέχρι το 2009, ομάδες όπως οι Spurs, Rockets και Thunder ενσωμάτωσαν το σύστημα στις προπονήσεις και στους αγώνες. Το 2009, οι Orlando Magic το χρησιμοποίησαν στους τελικούς NBA, σηματοδοτώντας την επίσημη αποδοχή του από τη λίγκα.

Η αξιοποίηση των big data επιτρέπει στους προπονητές να καθοδηγούν, να επιλέγουν και να εκπαιδεύουν καλύτερα τους παίκτες, ενώ οι πληροφορίες διαμοιράζονται και με τα μέσα ενημέρωσης για εμπορικά οφέλη.

Η Περίπτωση των Golden State Warriors

Οι Warriors χρησιμοποίησαν την ανάλυση δεδομένων για να αλλάξουν ριζικά την αγωνιστική τους εικόνα. Συνεργάστηκαν με εταιρίες ανάλυσης δεδομένων για να αξιολογήσουν στοιχεία από το SportVU και άλλες πλατφόρμες. Μία από τις πιο ενδεικτικές διαπιστώσεις ήταν πως όσο περισσότερο ο Stephen Curry επιχειρούσε

τρίποντα, τόσο αυξάνονταν οι πιθανότητες νίκης της ομάδας. Το αποτέλεσμα ήταν η κατάκτηση τριών πρωταθλημάτων μεταξύ 2015-2018.

Οι Clippers και το CourtVision

Οι Los Angeles Clippers δημιούργησαν το δικό τους σύστημα, CourtVision, που προσφέρει προβλέψεις και οπτικοποιήσεις αγώνων σε πραγματικό χρόνο. Η τεχνολογία αυτή επέτρεψε στους θεατές να κατανοήσουν βαθύτερα το παιχνίδι, ενώ συνέβαλε και στην αύξηση της αξίας του franchise τους. Από το 2015, η ομάδα επένδυσε σε τεχνολογικές καινοτομίες και πλέον παράγει έσοδα της τάξης των 2,6 δισεκατομμυρίων δολαρίων ετησίως. Σύμφωνα με τα δεδομένα του StatMuse (2022), ο δείκτης defensive rating παρέχει σαφή εικόνα της αμυντικής απόδοσης..

SportRadar και Στατιστική Υποδομή

Η SportRadar είναι άλλη μία σημαντική εταιρεία που συλλέγει δεδομένα για το NBA, προσφέροντας υπηρεσίες σε μέσα, στοιχηματικές εταιρείες και τις ίδιες τις ομάδες. Η συνεργασία της με το NBA είναι μακροχρόνια και περιλαμβάνει ακόμα και δημοσιογράφους δεδομένων που βρίσκονται στο γήπεδο κατά τη διάρκεια των αγώνων.

Βαθύτερα Μετρικά: PER, WS, TS%, Shot Charts

Σύμφωνα με στατιστικά δεδομένα παικτών από το Basketball Reference (Basketball-Reference, 2022), η αποδοτικότητα των αθλητών μπορεί να συγκριθεί σε βάθος χρόνου με βάση δείκτες όπως το PER και το Win Share.

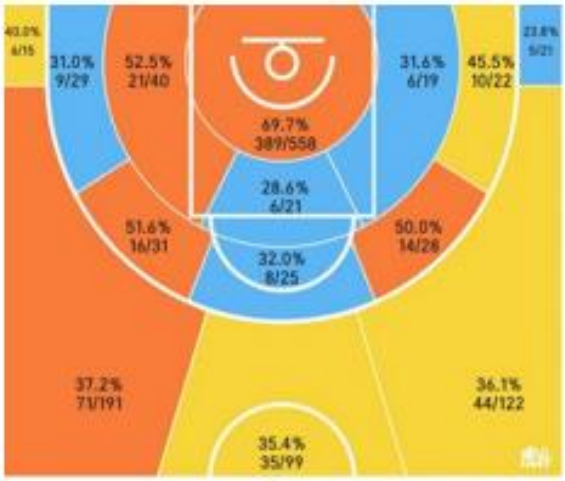
Το NBA χρησιμοποιεί προηγμένα στατιστικά μοντέλα:

Player Efficiency Rating (PER): Μετρά την παραγωγικότητα ανά λεπτό συμμετοχής. Παρότι χρήσιμο, έχει περιορισμούς στην αποτύπωση αμυντικών δεξιοτήτων.

Win Shares (WS): Αντιστοιχεί σε πόσες νίκες συμβάλλει ο κάθε παίκτης, χωριζόμενο σε επιθετικό και αμυντικό σκέλος. Έχει μεγαλύτερη σημασία για τη διοίκηση ομάδων.

True Shooting Percentage (TS%): Δείχνει την αποτελεσματικότητα σκοραρίσματος, συμπεριλαμβάνοντας δίποντα, τρίποντα και βολές.

Shot Charts: Χρησιμοποιούνται για τον σχεδιασμό τακτικών. Για παράδειγμα, οι Lakers προσαρμόζουν επιθέσεις του LeBron James με βάση τα σημεία όπου σουτάρει με μεγαλύτερη ευστοχία.



Εικόνα 16: Shooting Chart του LeBron James – 2022/2023 season

Μελλοντικές Εφαρμογές και Καινοτομίες

Η ανάλυση δεδομένων δεν περιορίζεται μόνο στις αγωνιστικές επιδόσεις. Μπορεί να εφαρμοστεί και στον τομέα της πρόληψης τραυματισμών μέσω αισθητήρων που παρακολουθούν φυσικά χαρακτηριστικά, φορτίο άσκησης και δείκτες πρόβλεψης κακώσεων, όπως εγκεφαλικές δονήσεις σε συνάρτηση με τον ΔΜΣ και τη μηχανική φόρτιση.

Η Σημασία της Ανάλυσης Δεδομένων στο Μπάσκετ και Πέρα από το NBA

Η τεχνολογία big data μεταμορφώνει το μπάσκετ. Παρέχει τη δυνατότητα λήψης τεκμηριωμένων αποφάσεων για επιλογές παικτών, στρατηγική αγώνων, εμπορικές συνεργασίες και πρόληψη τραυματισμών. Η χρήση τέτοιων δεδομένων ενισχύει την αποτελεσματικότητα των ομάδων και μειώνει το ρίσκο.

Το NBA έχει ηγετικό ρόλο σε αυτή την εξέλιξη, λειτουργώντας ως πρότυπο καινοτομίας για άλλες λίγκες και συλλόγους διεθνώς. Ο τρόπος με τον οποίο έχει ενσωματώσει την τεχνολογία, από το επίπεδο των στατιστικών μέχρι τις εμπορικές χρήσεις, δημιουργεί ένα νέο υπόδειγμα για το πώς μπορεί ο αθλητισμός να γίνει πιο στρατηγικός, έξυπνος και επικερδής (Deng & Zhong, 2020).

ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ

Η ανάλυση δεδομένων στον αθλητισμό έχει αναδειχθεί ως ένας από τους πλέον κρίσιμους παράγοντες για τη βελτιστοποίηση της απόδοσης, τη στρατηγική λήψη αποφάσεων και τη συνολική εξέλιξη των αθλητικών συλλόγων και των αθλητών. Μέσω της χρήσης στατιστικών μοντέλων και αλγορίθμων, οι ομάδες αποκτούν πρόσβαση σε πληροφορίες που προηγουμένως δεν ήταν διαθέσιμες ή ήταν βασισμένες αποκλειστικά στη διαίσθηση των προπονητών.

Τα μοντέλα και οι αλγόριθμοι, μέσω της μηχανικής μάθησης και της στατιστικής ανάλυσης, βοηθούν στην αξιολόγηση της απόδοσης παικτών, την πρόβλεψη αποτελεσμάτων, την ανίχνευση ρίσκου τραυματισμών και τη διαμόρφωση εξατομικευμένων προγραμμάτων προπόνησης. Παράλληλα, παρέχουν στρατηγική καθοδήγηση για αποφάσεις σχετικές με μεταγραφές, τακτικές και το αγωνιστικό πλάνο, ενώ η δυνατότητα ανάλυσης μεγάλων όγκων δεδομένων σε πραγματικό χρόνο επιτρέπει ταχύτερη και πιο τεκμηριωμένη προσαρμογή κατά τη διάρκεια αγώνων.

Οι ομάδες που εφαρμόζουν τέτοιες τεχνολογίες βελτιώνουν τόσο τη λειτουργία των εσωτερικών τους συστημάτων όσο και τα αποτελέσματά τους στον αγωνιστικό χώρο. Η αποτελεσματικότητα αυτή ενισχύει τη βιωσιμότητα των συλλόγων και επιτρέπει μακροπρόθεσμο σχεδιασμό βασισμένο σε αντικειμενικά δεδομένα.

Η σημασία της ανάλυσης δεδομένων για το μέλλον του αθλητισμού είναι θεμελιώδης. Σε έναν χώρο που γίνεται όλο και πιο ανταγωνιστικός, οι ομάδες που δεν υιοθετούν σύγχρονες μεθόδους ανάλυσης μένουν πίσω, τόσο σε επίπεδο απόδοσης όσο και στη διαχείριση των πόρων τους. Η σύγχρονη τάση δείχνει ότι η τεχνολογία και η επιστήμη δεδομένων είναι πλέον αναπόσπαστο κομμάτι του σύγχρονου αθλητισμού.

Τα case studies του Brentford FC και του NBA αποτελούν χαρακτηριστικά παραδείγματα του τρόπου με τον οποίο η ανάλυση δεδομένων μπορεί να μεταμορφώσει πλήρως την αγωνιστική και οικονομική πορεία ενός συλλόγου ή ενός οργανισμού. Η Brentford, με περιορισμένο προϋπολογισμό αλλά υψηλή επένδυση σε ανάλυση δεδομένων, κατάφερε να ανέβει στην Premier League και να σταθεροποιηθεί, ενώ ομάδες του NBA όπως οι Golden State Warriors χρησιμοποίησαν

προηγμένα συστήματα για να κατακτήσουν τίτλους και να βελτιστοποιήσουν το παιχνίδι τους σε κάθε επίπεδο.

Τα ευρήματα αυτά μπορούν να εφαρμοστούν και σε άλλα αθλήματα, όπως το τένις, ο στίβος και η κολύμβηση, όπου η ατομική απόδοση και οι βιομετρικές μετρήσεις μπορούν να αξιολογηθούν λεπτομερώς για τη δημιουργία εξατομικευμένων πλάνων προπόνησης, την αποφυγή τραυματισμών και τη μέγιστη αγωνιστική απόδοση.

Η γενικότερη συζήτηση που προκύπτει αφορά την ισορροπία ανάμεσα στην ανθρώπινη κρίση και τα δεδομένα: Αν και τα analytics παρέχουν ισχυρά εργαλεία, η σωστή ερμηνεία τους παραμένει ανθρώπινη υπόθεση. Επίσης, εγείρονται ηθικά και νομικά ζητήματα σχετικά με την προστασία προσωπικών δεδομένων και την υπερβολική εξάρτηση από την τεχνολογία.

Η ανάλυση δεδομένων δεν αποτελεί απλώς μια καινοτομία αλλά έναν πυλώνα εξέλιξης του σύγχρονου αθλητισμού. Οι ομάδες που κατανοούν, επενδύουν και ενσωματώνουν την επιστήμη δεδομένων, βρίσκονται σε σαφές ανταγωνιστικό πλεονέκτημα και θέτουν τα θεμέλια για μακροπρόθεσμη επιτυχία.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- **2021-22 NBA Player Stats: Advanced**. Basketball-Reference.Com. https://www.basketball-reference.com/leagues/NBA_2022_advanced.html. Accessed 5 Aug. 2022.

- Cai, L., Zhao, C., & Wang, X. (2021). *Situation and lessons of application of NBA big data technology*. 2021 International Conference on Information Technology and Contemporary Sports (TCS), pp. 228-231. doi: 10.1109/TCS52929.2021.00054

- Constantinou, A. C., & Fenton, N. E. (2013). *Determining the level of ability of football teams by dynamic ratings based on the relative discrepancies in scores between adversaries*. Journal of Quantitative Analysis in Sports, 9(1), 37-50.

- Deng, W., & Zhong, E. (2020). *Analysis and Prediction of Soccer Games: An Application to the Kaggle European Soccer Database*. Insight-Statistics, 3(1), 1-6.

- Haefner, J. *9 Stats That Every Serious Basketball Coach Should Track*. Retrieved 6 August 2022, from https://www.breakthroughbasketball.com/stats/9_stats_basketball_coach_should_track.html

- Hayhurst, C. (2020). *How NBA Data Analytics is Changing the Sport*. Retrieved 6 August 2022, from https://www.dell.com/en-us/perspectives/tech-takes-the-court-nba-teams-turn-to-data-and-analytics-for-a-competitive-edge/

- Hollinger's NBA Player Stats - ESPN Insider. ESPN.Com. http://www.espn.com/nba/hollinger/statistics. Accessed 5 Aug. 2022.

- Hvattum, L. M., & Arntzen, H. (2010). *Using ELO ratings for match result prediction in association football*. International Journal of Forecasting, 26(3), 460-470.

- Igiri, C. P. (2015). *Support vector machine—based prediction system for a football match result*. IOSR Journal of Computer Engineering (IOSR-JCE), 17(3), 21-26.

- Ken, B. (2015). *Warriors 'wearable' weapon? Devices to monitor players while on the court*. Retrieved 6 August 2022, from https://www.cbssports.com/nba/news/warriors-wearable-weapon-devices-to-monitor-players-while-on-the-court/

- Moneyball meets football | *Brentford's unique transfer model*. (Πηγή: Brentford FC)

- *Nylon Calculus: Reinventing PER*. FanSided, 31 Jan. 2017, https://fansided.com/2017/01/31/nylon-calculus-reinventing-per/

- Pezzana, C. (2021). *The NBA and Data Analytics*. Retrieved 6 August 2022, from https://www.linkedin.com/pulse/nba-data-analytics-carlo-pezzana
- *SportVU Basketball - Stats Perform*. Retrieved 6 August 2022, from https://www.statsperform.com/team-performance/basketball/optical-tracking/
- *Sports Analytics 101: An Intro to Advanced Stats*. Hudl Blog. (2016). Retrieved 6 August 2022, from https://www.hudl.com/blog/sports-analytics-101-an-intro-to-advanced-stats
- StatMuse. *Allen Iverson Defensive Rating | StatMuse*. https://www.statmuse.com/nba/ask/allen-iverson-defensive-rating. Accessed 5 Aug. 2022.
- Tavares, A. (2018). *Predicting Results of Brazilian Soccer League Matches*. University of Wisconsin-Madison.
- Tilp, M., & Schrapf, N. (2015). *Analysis of tactical defensive behavior in team handball by means of artificial neural networks*. IFAC-PapersOnLine, 48(1).
- Van Haaren, J., & Van den Broeck, G. (2015). *Relational learning for football-related predictions*. In *Latest advances in inductive logic programming* (pp. 237–244).
- Vanschoren, J., Blockeel, H., Pfahringer, B., & Holmes, G. (2012). *Experiment databases*. Machine Learning, 87(2), 127–158.
- Vaz, M. S., Ribeiro, Y. S., Pinheiro, E. S., & Del Vecchio, F. B. (2019). *Psychophysiological profile and prediction equations for technical performance of football players*. Revista Brasileira de Ciências do Esporte, 41, 215–221.
- *What Football Analytics Can Teach Successful Organisations | Rasmus Ankersen | TEDxManchester*
- *Brentford's Moneyball Strategy Explained*. Reddit, r/PremierLeague.