26

Artificial Intelligence and Genealogy

Nicoladie Tam, Ph.D.
University of North Texas
Denton, Texas, 76203, USA

Abstract

This review explores quantitative genealogical and forensic analyses of genetic data, illustrated with case examples that introduce the core principles of probabilistic inheritance and analytical techniques for determining ancestral lineages and identity matches. The discussion emphasizes the use of short DNA polymorphism variant sequences, including single nucleotide polymorphisms (SNPs), short tandem repeats (STRs), and microhaplotypes (MHs), as biomarkers. Forensic DNA index system databases were discussed

The illustrative case examples include tracing the ancestral origins of domesticated dogs from wolf populations by applying principles of inheritance and probabilistic analyses to identify their likely dual-ancestry ancestors. Another example tracks common ancestors before the branching of the DRD4-7R dopamine D_4 receptor gene variant tree, characterized by the 7-repeat variant VNTR (variable number tandem repeats), serving as a biomarker to analyze the natural selection process that led to beneficial outcomes despite the adverse effects of disorders like ADHD (attention deficit hyperactivity disorder) and novelty-seeking behavior. Additionally, another case example discusses tools that address the limitations of DNA samples in non-invasive prenatal paternity testing (NIPPT) by utilizing maternal blood samples to obtain fetal DNA.

The alternative artificial intelligence (AI) approach in genealogical analysis was discussed, focusing on two distinct AI methods for addressing complex problems. The traditional algorithmic AI approach relies on programming analytical methodologies into expert systems with clearly defined statistical inference techniques to solve these problems analytically. However, it struggles to address unexpected scenarios and cannot manage gaps in missing genetic data.

The alternate branch of the AI approach addresses these issues by employing machine learning to derive solutions without specifying the exact methodologies or

algorithms needed to tackle these problems. Case studies illustrate the principles and history of neural computation alongside machine learning, demystifying the AI magic that excels at summarizing large datasets while lacking explainable reasoning in its generated results. The advantages and practical limitations of machine learning provide informed criteria for evaluating suitable AI model performance.

The limitations include the requirement for a large dataset for training to identify suitable solutions and learn from examples to summarize data through input-output correlation, which is a time-consuming process. Depending on the AI models using supervised or unsupervised learning paradigms, they may or may not need human guidance during training with known solutions. These tools can help fill gaps in genealogical databases, and while these outcomes are possible, they may not always have a rational explanation. Their solutions may not lead to the optimal one, as other alternatives could be more effective.

Until future AI models improve their reasoning abilities, it is wise to exercise caution when concluding. Nonetheless, these systems effectively automate the organization of genealogical databases from written and spoken records, and they can produce automated responses to users' inquiries by utilizing natural language processing to address questions from everyday users.

Table of Contents

26
Artificial Intelligence and Genealogy

26.1 INTRODUCTION

Our uniqueness arises from genetic variations that manifest as distinct phenotypic traits. Environmental factors, including epigenetics, also impact gene expression, resulting in diverse phenotypes (or characteristics) even among individuals with the same genotype (or gene sequences), such as identical twins who vary due to environmental influences. The interactions between environments and genes create variations among us that contribute to our uniqueness. Without these variations, we would resemble factory-produced robots, distinguishable only by their serial numbers.

26.1.1 AN OVERVIEW OF THE QUANTITATIVE ASSESSMENT OF GENETIC DATA

The methodology for identifying an individual's or population's ancestral lineage or inherited traits varies based on project objectives, particularly in forensic and genealogical studies. This overview discusses methods for quantitatively assessing genetic information and includes case examples to illustrate the outcomes of the applied analytical techniques. The theoretical principles of these methods are presented to facilitate an objective evaluation of the results. Additionally, the advantages and disadvantages of each technique are examined to support informed decision-making. Data availability, technology options, and performance outcomes can influence the selection of analytical tools and the necessity of definitive conclusions for the project's goals.

26.1.1.1 Rationale for Choosing Specific Biometric Markers

Evolution offers evidence for tracing generations through inherited patterns in gene sequences. These variations assist in identifying genealogical lineages and reveal the divergent branches that enhance our understanding of identity and ancestry. The gene sequences provide quantitative evidence of inherited patterns transmitted from prior generations.

26.1.1.2 The Use of Genetic Codes as Identifying Markers

Analyzing gene patterns can reveal previously unknown genealogical lineages based on inherited traits in family trees. Constructing a genealogical tree involves connecting members according to the likelihood of these patterns being passed down.

This process infers a descendant's inheritance from the similarities and differences in gene sequences, organizing the tree by positioning members based on the most probable inherited traits. The gene code pattern can be a fingerprint identifier for matching an individual's identity in forensic determinations.

26.1.1.3 The Human Genome as an Identifying Blueprint

Deoxyribonucleic acid (DNA) encodes genetic information through its molecular sequence found within the genes of all living organisms, from bacteria to humans. DNA provides the genetic instructions for protein synthesis, cell replication, and repair. It also influences the developmental timing of plants and animals, affecting their resilience and survival.

DNA serves as the genetic code found in a cell's chromosomes. The complete set of chromosomes is referred to as the genome. The human genome consists of 6 billion nucleotides organized into 23 pairs of chromosomes. The number of chromosomes varies across species; for instance, some butterflies possess over 400, whereas certain single-celled protozoa have 1600.

26.1.1.4 The Complete DNA Codes in the Human Genome

Nucleotides form long polymer chains that create the backbone of a double-stranded helix composed of deoxyribose and phosphate groups. Strong covalent bonds link the nucleotides, while weaker hydrogen bonds connect the strands, resulting in a double helix with complementary base pairs. Each strand pairs A with T and G with C at corresponding loci on the chromosomes. The complementary base pairs enable one strand to predict its counterpart, resulting in 3 billion unique nucleotide codes that reflect individual similarities and differences.

26.1.1.5 The Complexity of Genetic Codes

Analyzing three billion DNA codes in the human genome is a daunting task, whether the goals are to identify individuals for forensics, trace ancestral origins for genealogy, or connect genetic factors associated with disorders for medical research. Examining the similarities and differences can provide insights into our genetic heritage.

26.1.2 GENOME-WIDE ASSOCIATION STUDIES (GWAS)

Although the human genome sample size is substantial, various high-throughput techniques allow for the systematic analysis of DNA sequence variations across populations in genome-wide association studies (GWAS). These techniques identify genomic variants statistically linked to specific traits or the risk of a disease, providing insights into the genome's coding and non-coding regions [1,2].

26.1.2.1 The Use of Gene Variants to Minimize Size for Comparison

Human nucleotide sequence variations range from 0.1% to 0.4%. While DNA sequences share over 99% similarity, the combinatorial variations remain significant for the comparative analysis of three billion-letter codes. Fortunately, most variations consist of single nucleotide substitutions or repetitive patterns. Comparing short DNA segments with variations in a population is sufficient for tracing or identification. These short sequences allow for comparison without needing to sequence the entire genome.

26.1.3 DISTINCT SHORT DNA SEQUENCES AS BIOMETRIC IDENTIFIERS FOR TRACKING

Even though each individual's unique genome can serve as a biometric identifier that distinguishes one person from another, it is commonly called a biomarker when a shorter sequence is used. Variations in these gene biomarkers can verify individual similarities and differences, including susceptibility to disorders, behavioral outcomes, and forensic tracing.

26.1.3.1 The Application of Gene Variants in Medical and Genealogical Analyses

These gene variants clearly distinguish individuals, irrespective of their functions or the traits they influence. They serve as objective biometric markers for tracing ancestry, locating relatives, confirming parenthood, verifying an individual's identity, unveiling genetic predispositions to disorders, and providing insights into heritage.

26.1.3.2 The DNA Sequences for Analysis

These DNA sequences are the foundation for all genetic coding information, including coding and non-coding regions in the chromosomes. This information can be used for forensic or genealogical analysis and to identify genetic links to medical disorders. While various forms of data, such as historical records and personal accounts, can provide valuable insights for forensic and genealogical tracing, this chapter focuses on leveraging DNA data for these purposes.

26.1.4 THE ROLE OF NON-CODING DNA IN FORENSIC ANALYSIS

Not all gene sequences encode or regulate gene expression; some are non-coding and unexpressed. Coding regions are essential for studying effects on growth, development, and medical disorders [3]. Since non-coding sequences are not expressed, they do not typically influence trait outcomes, which can affect analysis. For these reasons, non-coding regions are commonly used for forensic or genealogical analysis.

26.1.4.1 Effects of Non-Coding DNA on Disorders

Although non-coding regions are not transcribed, they are associated with specific phenotypic variants identified by the GWAS project. Single nucleotide polymorphisms in non-coding DNA can influence intron activity and regulatory elements, such as promoters and enhancers. These variants may impact DNA methylation, histone modifications, transcription factor affinity, alternative splicing, and mRNA stability. These factors suggest that they can alter traits even though they are not directly transcribed [4]. In forensic analysis, it was once assumed that using non-coding regions would not affect traits, but emerging evidence may not support this assumption.

26.1.4.2 Amplification of DNA Sequence for Analysis

The polymerase chain reaction (PCR) amplification method is widely used to replicate a specific DNA sequence for analysis. Sequencing a shorter segment is simpler than sequencing the entire genome. Several other amplification and sequencing techniques are available for extracting a target DNA sequence for analysis.

26.1.5 SHORT SEQUENCE VARIANTS AS BIOMARKERS FOR IDENTIFICATION

To reduce the sequence size for analysis, one can select a specific short sequence for comparison, provided that the sequence distinguishes differences among individuals as identifiable markers. These short sequences can serve as biomarkers for assessing the likelihood of inheritance within a lineage, thereby narrowing the scope of comparison. When these short nucleotide segments provide sufficient data to trace inherited traits, they can yield enough information to create a genetic profile for forensic analysis.

26.1.5.1 Improve Matching Accuracy with Multiple Biomarkers

False positives inevitably occur in any testing due to errors, regardless of accuracy. Utilizing short-sequence biomarkers can match genetic traits more efficiently than analyzing the entire genome; however, false positives persist. Nevertheless, the error rate can be significantly reduced by employing multiple biomarker sequences from different loci (locations on a gene) instead of relying on just one.

For example, with a 10% false positive error rate (1 in 10) per sequence tested, two sequences result in a 1% error chance (10% x 10% = 1 in 100). Testing three sequences reduces this to 0.1% (10% x 10% x 10% = 1 in 1000). With 13 sequences, the false positive rate drops to 0.1^{13}% or 1 in a trillion. Thus, most forensic analyses use multiple biomarker sequences at various loci to determine matches.

26.2 AN OVERVIEW OF TRADITIONAL STATISTICAL INFERENCE

Identifying genealogical lineage relies on methods that evaluate the probability of inheriting DNA sequences from ancestors or matching individuals through genetic profiles instead of depending on fingerprint analysis. Traditional approaches use statistical inferences to assess the likelihood of trait transmission, which aids in constructing genealogical trees or matching DNA biomarker sequences for forensic analyses.

26.2.1 FALSE POSITIVES AND STATISTICAL CONFIDENCE

These methodologies rely on the statistical probability of matching a target population or gene sample. However, the possibility of error always exists. To minimize the risk of false positives, statistical methods lower error rates by increasing the confidence intervals in the analysis. A conclusion can only be drawn when there is sufficient statistical probability to establish a genealogical tree or evaluate the likelihood of matching an individual to a target sample.

26.2.2 DEPENDENCY ON A SPECIFIC DATABASE

However, comparing genetic data requires access to a database of DNA sequences from populations potentially related to the individual. The location on the tree can be identified if the matching sequences can be traced back to the genetic profiles already present in the database. Tracing the location on the tree becomes increasingly difficult when related profiles are absent from the database.

26.3 AN OVERVIEW OF ARTIFICIAL INTELLIGENCE IN GENEALOGY

This overview of using artificial intelligence (AI) in genealogy summarizes the capabilities of AI models. It discusses the advantages and limitations of this technology before exploring the differences among various AI models that impact genealogical analysis. Subsequent sections will present a more detailed discussion, featuring a case history and examples to illustrate the principles used in AI, offering objective criteria for evaluating the performance of different AI models genealogy.

26.3.1 THE ALTERNATIVE AI APPROACH BEYOND STATISTICAL INFERENCES

In contrast to the traditional standard approach, artificial intelligence techniques can analyze and generalize genetic information more effectively than conventional methods. They provide potential solutions that uncover new insights into genealogical

inheritance and bridge data gaps through generalizations. Natural language processing AI models can automate the categorization of written genealogical records [5] while offering human-like interactions to address inquiries related to genealogy by generalizing genetic profile information from the training datasets.

26.3.2 ADDRESSING THE DATA GAPS

AI effectively addresses issues related to missing data. Current AI technology employs machine learning (ML) to integrate extensive datasets and train systems to establish connections between data points. Once trained, these connections generate outputs from input queries, even when specific data points are absent from the training set. The system can interpolate missing information by leveraging the connectivity among relevant data. In forensic genealogy, if the database lacks genetic profile information, AI can estimate likely tree locations, which traditional algorithms may struggle to accomplish.

26.3.3 GENERALIZING GENEALOGICAL PATTERNS

Other applications of AI models in genealogy assess relationships without depending on traditional statistical methods. Rather than explicitly programming analytical techniques into computer software, AI models generate results for input queries by generalizing patterns learned from the interactions between inputs and outputs in the training dataset.

26.3.3.1 Responding to Genealogical Inquiries Using Natural Language Processing

For instance, AI can generate natural responses to genealogical inquiries using a large language model (LLM), such as ChatGPT. It can automate user requests and provide answers without human involvement. Generative AI can create interactive conversations with users to address questions based on training with a vast dataset.

26.3.4 CATALOGING NATURAL LANGUAGE GENEALOGICAL RECORDS

Applying AI in genealogy can extract information from historical records, including newspaper articles and oral histories, and organize it into databases. This cataloging process is often labor-intensive; automating it would minimize the need for human involvement and speed up information retrieval. Users can benefit from organized genealogical information for future queries and can pose questions to receive more efficient answers.

26.3.5 MACHINE LEARNING (ML) FOR LEARNING FROM TRAINING DATASETS

An AI system develops internal models by learning input-output relationships through machine learning (ML) without relying on human-provided solutions or algorithms. ML automates this process by applying learning rules to establish connections, including direct and higher-order relationships among relevant factors. The system utilizes a training dataset to improve its internal representations through trial and error, using an error minimization methodology to achieve optimal outcomes. It requires billions of examples in the dataset and billions of iterations in the trial and error process to converge on a solution, effectively linking input queries to system outputs.

26.3.5.1 Biases in Training Databases

The model may unintentionally introduce bias by optimizing internal representations based on a particular training dataset. While large datasets are essential, biased outcomes can arise if counterexamples are excluded. Due to the limited genealogical databases, conclusions drawn from specific datasets might be flawed. Furthermore, outliers are often misrepresented during the generalization process.

26.3.5.2 Limitations of Artificial Intelligence

The process by which the AI system reaches its conclusions is often unclear. It generates outputs by learning correlations from extensive datasets. Consequently, AI models excel at generalizing key data to produce results. They do not apply logical reasoning or specific rationale to deduce answers, such as using physiological or genetic knowledge of inheritance or employing statistical principles to infer outcomes. Therefore, AI outputs typically lack explanations unless explainable AI (XAI) is used to clarify the reasons based on the analysis of statistical input-output patterns. Without logical inferences and reasoning rationales, the AI system's conclusions may be plausible yet lack rationality.

26.4 GENEALOGICAL ANALYSIS METHODOLOGIES

Methods for genealogical analysis rely on the interactions between genotypes and phenotypes. Variations in DNA sequences can lead to differences in genotypic and phenotypic traits. These genotypes influence our physical characteristics, closely linking them to phenotypic qualities. Such variations emerge through the mixing and matching processes during DNA replication and recombination. Analyzing these sequences can uncover the genetic influences on inherited traits, which may offer survival advantages or disadvantages. As we examine probabilities on a genealogical tree, these variations can affect lineage branching and extinction.

26.4.1 ANALYSIS OF DNA PATTERNS

As mentioned earlier in the introduction, genetic codes are represented by the four-letter symbols A, C, G, and T. These sequences correspond to their respective chromosome pairs, with A pairing with T and C pairing with G. The analytical methodology focuses on the patterns of similarities and differences among a series of DNA sequences.

26.4.1.1 The Degree of Similarity and Difference

In genealogical tracing, the degree of similarity can indicate a level of relatedness. Identifying lineage becomes straightforward when comparing identical and non-identical codes in a DNA sequence as measures of relatedness. However, exact and non-identical sequences do not necessarily correlate with similarities or dissimilarities or imply inheritance for the reasons explained below.

26.4.1.2 The Difference Defined by a Single Nucleotide Sequence

When a DNA sequence varies by one or two nucleotides, should it be considered similar or different compared to a group? Classification is relative to the population with which it is compared and is influenced by levels of variability. Similarities or differences can change depending on the subpopulation used for comparisons. Assigning an individual to a lineage tree depends on their similarities and variations compared to other groups. The analysis involves more than just genetic similarities and differences; it requires additional methods to trace lineages based on how descendants inherit genes and survive.

26.4.2 PRINCIPLES OF INHERITANCE

Various factors influence inheritance patterns in sexual reproduction. This reproductive method accelerates the evolutionary rate by mixing recombinant DNA more effectively than random mutations, thereby enhancing survival. Consequently, the recombinant DNA sequences differ from those of either parent, while descendants inherit some traits but not all. Understanding these processes could assist in identifying suitable analytical tools for genealogical analysis.

26.4.2.1 Inheritance via Asexual Reproduction

In asexual reproduction through cloning from a single parent, the offspring's DNA is identical to that of the parent, as they are exact replicas of the organism. When cloned, the progeny of asexual reproduction inherit 100% of the parent’s traits, assuming there are no mutations or replication errors during reproduction. The process of inheritance in asexual reproduction is straightforward and certain. However, inheritance in sexual reproduction, which involves recombinant DNA, is more complex and probabilistic rather than certain.

26.4.2.2 Inheritance via Sexual Reproduction

Sexual reproduction combines DNA from both parents, producing offspring that inherit half of their traits from each. It accelerates evolution by mixing and matching in a trial-and-error recombinant reproductive process. Unlike asexual reproduction, the inheritance patterns in sexual reproduction are not definitive; they are probabilistic and depend on numerous factors that affect genealogical analysis.

26.4.2.3 Single-Stranded Chromosomes in Haploid Cells

During sexual reproduction, meiosis is a type of cell division that separates chromosomes into unpaired forms, resulting in haploid gametes, specifically sperm and egg cells. Each gamete contains one copy of each chromosome, made up of a single DNA strand. The chromosomes remain unpaired until fertilization when the sperm and egg fuse together.

26.4.3 DOUBLE-STRANDED CHROMOSOMES IN DIPLOID CELLS

During fertilization, a sperm cell inserts its nucleus into an egg, recombining chromosomes into paired strands. One unpaired chromosome is inherited from the mother and another from the father, resulting in a diploid cell. The fertilization process recombines unpaired DNA strands, giving the embryo half of the traits from each parent, while the chance of inheriting a specific strand remains probabilistic.

26.4.3.1 The Randomness of Producing a Male or Female Offspring

Sexual reproduction demonstrates the randomness of inheriting specific sex chromosomes. Parents can produce offspring with either XX chromosomes (indicating females) or XY chromosomes (indicating males) (see Figure 26.1). This process occurs by chance during fertilization. The likelihood of having a boy or a girl is a probabilistic phenomenon rather than an absolute certainty.

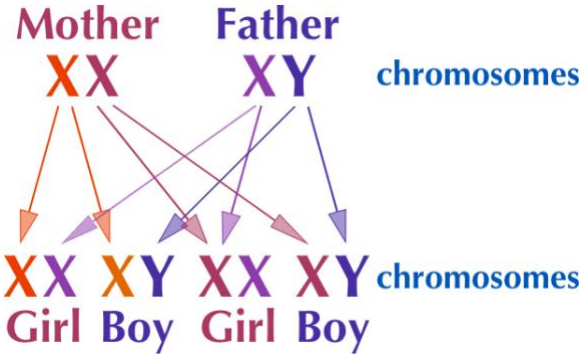

Figure 26.1. The diagram illustrates the recombinant X and Y chromosomes, showing a 50% (2 out of 4) probability of inheriting either an XX chromosome for a girl or an XY chromosome for a boy.

26.4.3.2 The Probability of Having a Girl or a Boy

The probability of inheriting a particular sex chromosome, such as the XX chromosome, is 50% (2 out of 4) for girls and 50% (2 out of 4) for boys (see Figure 26.2), assuming the fertilization process is not manipulated. It indicates the chance of having a girl or a boy is 50/50. This example illustrates the random chance involved in inheriting traits through sexual reproduction.

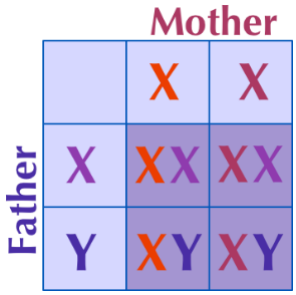

Figure 26.2. The inheritance table illustrates the recombinant XX and XY chromosomes from the mother and the father, similar to the inheritance shown in Figure **26.1**.

26.4.4 THE INFLUENCE OF SURVIVAL RATES ON PREVALENCE STATISTICS

Due to survival rates, population prevalence statistics can overestimate or underestimate inheritance probabilities. For example, females tend to live longer, leading to a male-to-female ratio of 49% to 51%. Prevalence data may misrepresent the frequency of gene variants based on survival rates: a male-linked variant shows a bias toward 49%, while a female-linked variant exhibits a bias toward 51%. This example

illustrates how survival hazard functions influence age-specific incidence statistics. Therefore, prevalence statistics may not accurately reflect occurrences of gene variants unless age-specific survival rates are considered.

26.4.4.1 The Hazard Function of Survival Rate for Each Age Group

The hazard function represents the cumulative distribution until death occurs. If a gene variant leads to early childhood deaths, those cases will not be counted if the deaths occurred before the population survey. Furthermore, if parents do not have children, the chances of passing down genes are nonexistent, not due to death but rather a lack of reproduction. The prevalent statistics can adjust for the survival rate within each age group for a more accurate representation.

26.4.4.2 Factors Influencing Parental Inheritance

Although all offspring's genes can be traced back to their parents, one-half of the parent's chromosomes are absent in the offspring. When a gene variant is used to track inheritance, that variant may or may not be present in the descendent's genes. The offspring might inherit the normal gene from the heterozygous carrier parent and receive a regular copy instead of the variant.

Forensic and genealogical tracing requires an analysis of how genes are passed down from each parent, which is crucial for drawing reliable conclusions. The probability of passing a gene variant to offspring depends on whether one or both parents carry it and whether they are heterozygous or homozygous carriers.

26.4.5 ESTIMATING THE LIKELIHOOD OF INHERITING A GENE VARIANT

Unlike asexual reproduction, the probability of passing this variant to the next generation is generally considered to be 50% for an offspring inheriting one chromosome from each parent. This scenario assumes that one parent is normal. In contrast, the other parent carries the variant and is heterozygous for that gene variant, meaning one chromosome copy has the variant. In contrast, the other is normal, similar to the rest of the population.

26.4.5.1 Inheritance Patterns as a Random Recombinant Process

Figure 26.3 illustrates how inheritance patterns vary depending on whether one or both parents are homozygous or heterozygous for the gene variant. The corresponding figure panel displays the likelihood of inheriting a specific gene variant. There are four possible combinations for transmitting a gene variant from the parents. One or both parents may be carriers of the variant, either as heterozygous or homozygous carriers. Figure 26.3 shows the associated probability of inheritance.

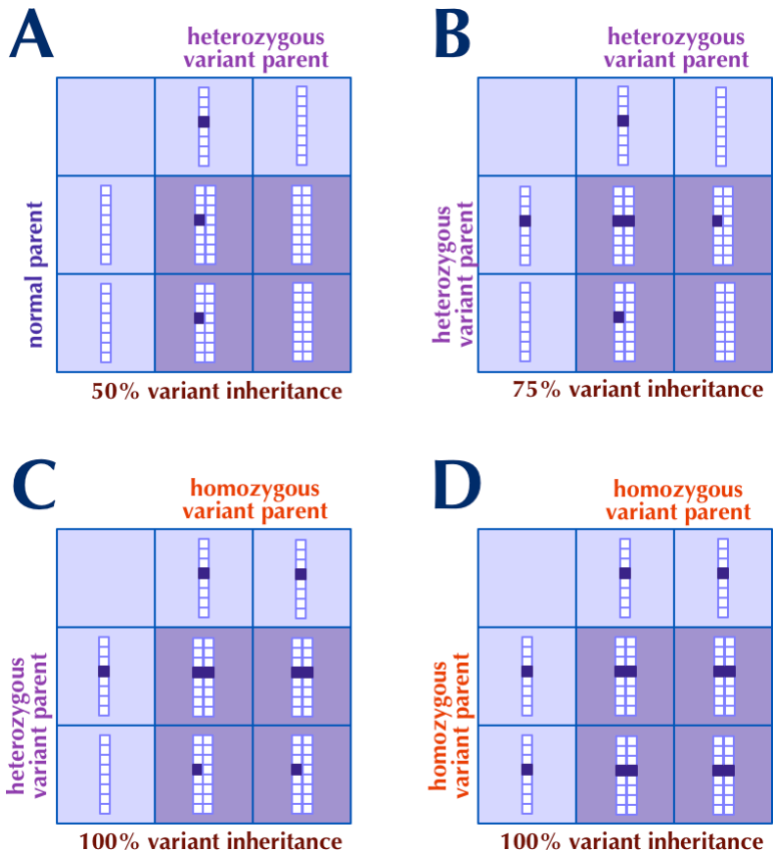

Figure 26.3. The table illustrates the inheritance patterns of gene variants passed on to the next generation. The probability of inheritance depends on whether one or both parents are heterozygous or homozygous. (The gene variant is indicated by a blue dot in the DNA sequence, while white dots represent the normal gene sequence.)

26.4.5.2 The Probability of Inheriting from a Heterozygous Parent and a Normal Parent

If one parent is a heterozygous carrier with one copy of the gene variant and the other is normal, there is a 50% chance that the offspring will inherit the variant (see Figure 26.3A). The offspring have a 50% chance of inheriting a heterozygous gene variant.

26.4.5.3 The Probability of Inheriting from Two Heterozygous Parents

The likelihood of inheritance rises to 75% when both parents are heterozygous (see Figure 26.3B). It suggests there is a 25% (1 in 4) chance of inheriting the variant from both parents, a 50% (2 in 4) chance of inheriting it from either parent, and a 25% (1 in 4) chance of not inheriting it. Therefore, the average probability remains 75% (3 in 4).

26.4.5.4 The Probability of Inheriting from Homozygous Parents

If one or both parents are homozygous, the likelihood increases to 100% (see Figure 26.3C&D). The offspring will undoubtedly inherit it with complete certainty. Therefore, the inheritance pattern varies depending on whether the parents are heterozygous or homozygous.

26.4.5.5 The Probability of Inheritance for Parents without Children

The preference for these variant traits in mate selection could significantly increase prevalence statistics, provided that the mating pairs have offspring. The likelihood of passing these variants to the next generation is nonexistent if they do not have children. This example illustrates how mate selection and reproductive success affect population prevalence statistics beyond random chance.

26.4.6 HARDY-WEINBERG EQUILIBRIUM

Genetic variation within a population will remain stable across generations if the reproductive process is entirely random. This phenomenon is referred to as Hardy-Weinberg equilibrium. It assumes random mating and reproduction within a large population free from disruptive influences. As a result, genotype and allele frequencies will remain unchanged.

26.4.6.1 Factors Influencing Hardy-Weinberg Equilibrium

However, real-world factors such as mutations, meiotic drive, natural selection, non-random mate selection, genetic drift, gene flow, and harmful alleles can disrupt this equilibrium. These factors may influence the estimated probability, especially when a conditional probability depends on a prior probability. Conversely, when tracing ancestral history, an offspring inherits one copy of each pair of gene sequences directly from each parent with certainty. Thus, the probability of inheritance relies on the direction of tracing and the analysis performed.

26.5 NUCLEOTIDE VARIANTS AS BIOMARKERS

Forensic analysis often emphasizes potential lineage over the effects of variations in gene expression. It typically selects non-coding regions that are not expected to be transcribed for polymorphism analysis. These short sequences act as biomarkers for identification rather than provide a comparison of the entire genome. A single nucleotide base in a DNA sequence may change due to mutations, replication errors, inheritance, or other factors. Nevertheless, there is at least a 50% chance that it will be passed on to the next generation since offspring inherit one chromosome from each parent.

26.5.1 SINGLE NUCLEOTIDE POLYMORPHISM

Single nucleotide polymorphisms (SNPs) are variations of a single nucleotide within an identical DNA sequence (see Figure 26.4). Polymorphism refers to the substitution of one nucleotide for another within a sequence. It can arise from mutations, replication errors, or other initial causes. The substituted nucleotide is then passed down to subsequent generations, leading to variations among similar individuals in a population.

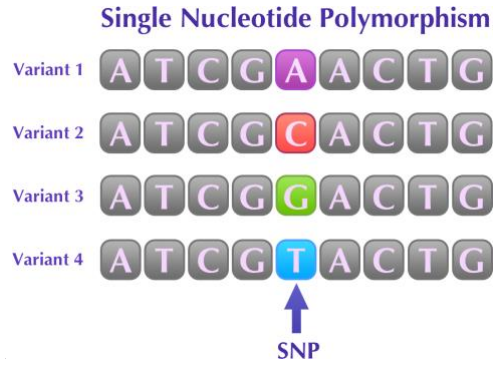

Figure 26.4. A single nucleotide polymorphism (SNP) in a single-stranded DNA sequence represents a variation at a specific nucleotide position within an otherwise identical DNA sequence, differing solely at that position. There are four possible variants of an SNP, each corresponding to substituting one of the four different nucleotide bases, assuming that deletion or insertion is excluded.

26.5.1.1 Homozygous and Heterozygous Alleles

An allele is one of two or more variations of a DNA sequence (a single base or a segment of bases) found at a specific genomic location. An individual inherits two alleles – one from each parent (see Figure 26.5). If the two alleles are identical, the individual is homozygous for that allele; if they are different, the individual is heterozygous.

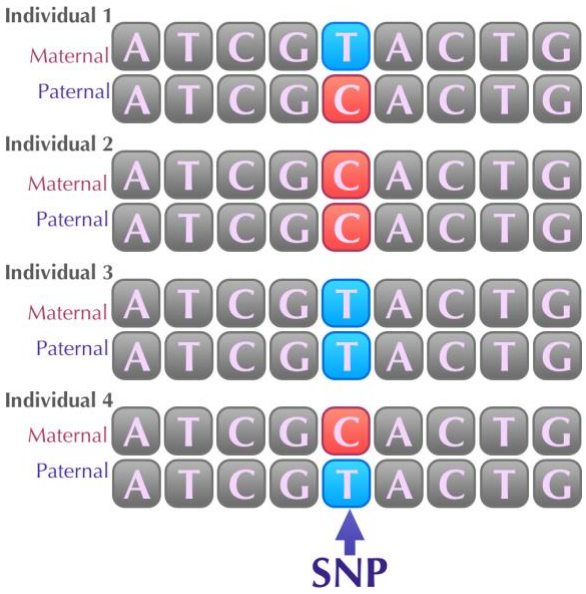

Figure 26.5. A single nucleotide polymorphism (SNP) in a double-stranded DNA sequence. One strand is inherited from the mother, while the other comes from the father. This diagram illustrates a variation in a single nucleotide within an otherwise identical DNA sequence across four individuals.

26.5.2 POLYMORPHIC SEQUENCES AS BIOMARKERS

The biomarker sequence is frequently chosen for its polymorphic characteristics, which exhibit nucleotide variations in gene sequences among individuals within a population. Common biomarkers include single nucleotide polymorphisms, short tandem repeats, and microhaplotypes. They are the most prevalent polymorphisms.

26.5.2.1 Polymorphism Tracing

If the nucleotide alteration originated from a mutation in a parent, then the descendant's inheritance is unique to that parent. However, if it was passed down from previous generations, the likelihood of inheriting the polymorphism depends on its statistical incidence in the population. This probability can be estimated by comparing it to the occurrence in the target population.

26.5.3 SINGLE NUCLEOTIDE POLYMORPHISM AS A BIOMARKER

When inherited, a single nucleotide difference can serve as a unique biomarker to trace a direct lineage from a parent instead of comparing complete genomic sequences. This similarity arises because the sequence matches the entire population, differing by only a single nucleotide passed down from that ancestor during ancestry tracing. Even a

single nucleotide variation can distinctly identify an individual based on these DNA variants.

26.5.3.1 Comparison with a Target Population

Suppose a single base substitution differentiates an individual's DNA sequence from the general population. In that case, it may be inherited from the mother, father, or both parents since each chromosome strand originates from one parent. This variation aids in identifying descendants from a specific subpopulation and provides insights into their ancestral characteristics. Additionally, it can help identify individuals based on inherited patterns. For analysis, the likelihood of inheriting a particular polymorphism is compared to its frequency in the target population [6].

26.5.4 CASE EXAMPLE: TRACING THE ANCESTRY OF DOMESTICATED DOGS

The methodologies for tracing the ancestry of dog domestication are illustrated through a case study. While it is widely accepted that gray wolves are the ancestors of dogs, there is no consensus on when, where, and how this transition took place. This case study will demonstrate the analytical principles used to trace the ancestral origin.

26.5.4.1 Tracing the Distribution of Ancient Wolves Across Time and Space

Archaeological evidence from skeletal remains indicates that modern dogs first appeared approximately 14,000 years ago. Genetic data shows their divergence from wolves occurred between 40,000 and 14,000 years ago. Identifying genetic diversity in wolves over time and across regions may clarify which populations were most closely related to the ancestors of dogs.

26.5.4.2 Tracing Ancestral Origins Using SNP Biomarkers

SNP genotypes were collected from a dataset covering the past 100,000 years, including data from contemporary wolves, ancient dogs, and various canids. An ancient dhole genome served as a control outgroup [7]. The analysis indicated that 69% of the wolves were male, based on X-chromosome DNA sequencing from 66 genomes across Europe, Siberia, and northwestern North America. This finding corresponds with a similar overrepresentation of males observed in the genomes of ancient woolly mammoths, bison, brown bears, and domestic dogs. If ancient dogs had different affinities with wolves before domestication, then gene flow from dogs could not have impacted wolves.

26.5.5 PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) uncovers hidden factors in datasets by identifying essential variables. This versatile tool is applicable to various data types, revealing latent variables as linear combinations of the original ones. PCA identifies the main contributing factors by rotating the principal component axes to minimize deviation. This iterative process continues until the remaining variables contribute negligibly. The first few components capture the most variance in descending order, thereby reducing dimensionality by excluding variables that contribute minimally. The new variables simplify analysis without compromising data fidelity. Although abstract variables capture the essence of the data, they may lack direct physical interpretation unless connected to recognizable phenomena such as thought or consciousness. These variables emerge from quantifiable neural processing that PCA reveals.

26.5.5.1 Data Analysis of Unidentified Latent Factors

PCA is an analytical method for exploratory analysis that identifies latent factors influencing results. This technique provides unbiased insights into datasets. If unsure which factors affected your data, PCA can reveal them by analyzing variables through new combinations to highlight significant ones. Even without prior knowledge of the factors, this method uncovers them using a best-fit model, yielding insightful outcomes.

26.5.5.2 Case Example: Abstract Factors Revealed by PCA

PCA creates abstract factors from multiple variables, streamlining complex data into more straightforward representations. For instance, census data on income, education, rent, car value, and postal code may seem unrelated. The data exists in a five-dimensional space with five variables, making visualization challenging. PCA addresses this by generating a single factor that combines all five variables into one, positioning data points along an axis that best fits. It identifies an abstract “socioeconomic” axis that aligns with the data, effectively reducing the five variables to one and clarifying the impact of a single factor on society more effectively than the original variables did.

26.5.5.3 The Versatility of Principal Component Analysis

Unlike AI tools that often lack explanations, PCA uses statistical models to represent data across new dimensions of contributing variables. This versatile technique applies to most data types. It offers statistical transparency, reliability, and reproducibility, even for those with limited statistical knowledge. It uncovers unknown factors without prior knowledge of the contributing variables that govern the data.

26.5.5.4 Factors Identified by PCA

The major factors are the principal components that account for most of the overall variance, while the remaining factors are insignificant as they contribute little to that variance. Mapping the data into the principal components reduces dimensionality while preserving the essence of the data. When PCA extracts features using a variable, the principal component axes represent the feature vectors. When it describes the variations in the data through an equation, that variable becomes a feature, with the feature vector pointing toward the principal component axes.

26.5.6 PEARSON CORRELATION FOR IDENTIFYING POPULATION CLUSTERS

When PCA is applied to a matrix representing shared genetic drift, it indicates that ancient wolves cluster by age rather than geography. This finding is corroborated by the Pearson correlation coefficient, which measures the linear relationship between datasets. The Pearson correlation assesses the similarity between two datasets by comparing their attributes, producing a score that ranges from -1 to +1. A high score signifies strong similarity, while a score close to zero indicates no correlation.

26.5.7 GENE FLOW DIRECTIONALITY ANALYSIS

The analysis of gene flow directionality indicated that Siberian ancestry spread to Europe via the Bering land bridge from Alaska and the Yukon around 10,000 years ago, but not vice versa. This allows for the deduction of ancient wolf migration patterns.

26.5.8 MATERNAL LINEAGE ANALYSIS USING MITOCHONDRIAL DNA

Mitochondrial DNA is inherited solely from mothers, making it a valuable marker for tracing maternal lineage. Analyzing maternal mitochondrial DNA inheritance sheds light on the ancestry of wolves as they migrated from Siberia to Europe. Maternal inheritance acts as a powerful tool for defining inheritance patterns.

26.5.9 ANALYSIS OF ALLELE PREVALENCE IN NATURAL SELECTION

Natural selection was confirmed by analyzing allele preferences in a dataset spanning approximately 100,000 years (around 30,000 generations). The study assessed each variant regarding allele frequency and timeframe across 72 ancient and 68 modern wolves within 24 genomic regions while accounting for genetic drift to reduce false positives. The survival of Pleistocene wolves resulted in rapid adaptations in selected alleles.

26.5.10 DUAL ANCESTRY MODEL

PCA rejected a single ancestor for Near Eastern dogs. A wolf admixture model similarly confirmed varying ancestry proportions that accounted for asymmetries. The dual ancestry model indicates that Arctic dogs experienced less Western influence, while the Western components in Near Eastern and African dogs are associated with Near Eastern wolves.

26.5.10.1 Dual Origins of Ancestral Trees

The conclusion suggests two independent domestication events: one from Eastern and Western ancestors and another from a separate occurrence in the East, accompanied by Western admixture across multiple lineages [7]. Therefore, the assumption that a single origin must be identified should be cautiously approached in genealogical tree analysis.

26.5.11 CONSIDERATIONS IN HYPOTHESIS TESTING

The above example illustrates the importance of using appropriate analytical techniques to draw valid conclusions. Misleading results can occur without the proper analytical tools or assessments to address underlying assumptions.

If the hypothesis is based on a single lineage assumption, it may overlook potential evidence of dual ancestry. Dog domestication likely happened independently among various geographical populations due to their ability to respond to commands and assist in hunting during the evolution of human hunter-gatherers.

26.5.11.1 Self-fulfilling Prophecy

Proving a hypothesis with supportive evidence is insufficient for concluding and requires analysis that challenges the hypothesis. A hypothesis is only valid if it cannot be disproven. Proving a conjecture without disproof can lead to a self-fulfilling prophecy. Proving a conjecture without disproof reflects a subconscious bias when the proof focuses solely on validating evidence without considering its potential invalidity.

26.5.11.2 Challenging the Validity of Assumptions

This caveat often occurs when conclusions depend on preconceived notions assumed to be validated before any analysis. Consequently, the analysis selectively incorporates only supporting evidence while disregarding contradictory evidence. These represent subconscious biases in any analysis that one should be aware of.

26.5.11.3 Validating Hypothesis Testing Without Refuting It

The common fallacy in hypothesis testing is presenting only evidence that supports the hypothesis while failing to disprove it. A hypothesis is considered proven only if it cannot be refuted. Confirming its validity without showing that it might be incorrect reinforces a self-fulfilling prophecy.

26.5.11.4 Including an Opposing Hypothesis in the Analysis

Proposing an alternative hypothesis is essential to avoiding premature conclusions. A robust hypothesis must be backed by strong evidence and a counterargument. An alternative hypothesis helps to dispel myths and offers a possible explanation if the evidence does not support the original hypothesis.

26.6 CONSIDERATIONS FOR LIMITED DNA QUALITY AND QUANTITY

Utilizing a short DNA segment for forensic analysis is particularly crucial when DNA samples are of low quality or limited quantity, as this ensures accurate human identification or paternity testing. Only a small amount of fetal DNA is found in maternal blood samples without performing an amniocentesis for prenatal paternity testing. In forensic analysis at crime scenes, a small amount of degraded DNA may be available. Consequently, specialized techniques will be essential to perform DNA analysis on limited sample sizes and quality.

26.6.1 CASE EXAMPLE: NON-INVASIVE PRENATAL PATERNITY TESTING

Conventional prenatal paternity testing involves invasive amniocentesis, which carries a risk of miscarriage during the collection of amniotic fluid or umbilical cord blood. An alternative method is to obtain fetal DNA samples from maternal blood, though the quantity is limited. Fortunately, techniques are available to analyze insufficient fetal DNA for conclusive paternity identification. Non-invasive prenatal paternity testing (NIPPT) collects small amounts of fetal DNA from maternal blood for analysis, thereby eliminating the risks associated with amniocentesis.

26.6.1.1 Limited Quantity of Fetal Fraction DNA in Maternal Blood Samples

Fetal fraction refers to the proportion of cell-free fetal DNA present in maternal plasma. It varies among individuals and is influenced by gestational age and weight. It is about 15% on average, ranging from less than 4% to over 30%. This parameter is crucial for the accuracy of cfDNA-based prenatal paternity tests [8].

26.6.1.2 Maternal Cell-Free Fetal DNA Analysis

Maternal cfDNA paternity testing typically utilizes STR or SNP genotyping. The increased polymorphism of STRs benefits forensic Combined Paternity Index (CPI) and Cumulative Probability of Exclusion (CPE) analyses. However, cfDNA sizes ranging from 140 to 160 bp may exceed STR coverage, requiring additional techniques.

26.6.1.3 Improved Analytical Models for NIPPT

Models that incorporate fetal fraction and genotype probability improve the evaluation process. The Poisson-based fetal fraction model offers accurate estimates of fetal fractions without depending on known biological parents. Meanwhile, the genotype probability model enhances statistical power by merging the Poisson distribution with sequencing error rates. New CPI and CPE models provide more precise estimates based on NIPPT characteristics. Additionally, a T-test model can identify sample contamination due to its sensitivity to abnormal data [8].

26.7 VARIABLE NUMBER TANDEM REPEATS AS BIOMARKERS

In addition to single nucleotide polymorphisms, other types of DNA variants exist as repeated patterns in succession. These variants often comprise short sequences of repeats on the chromosome that vary from one individual to another. The number of tandem repeats can differ within the general population.

26.7.1 VARIABLE NUMBER TANDEM REPEATS

Short tandem repeats (STRs) are short DNA sequences made up of tandemly repeated units of 1 to 6 base pairs (bp), leading to sequences that can differ in length by up to 100 nucleotides (nt) (see Figure 26.6). Due to the variation in the number of repetitions among individuals, they are also known as variable number tandem repeats (VNTRs). Furthermore, they are also called microsatellites or simple sequence repeats.

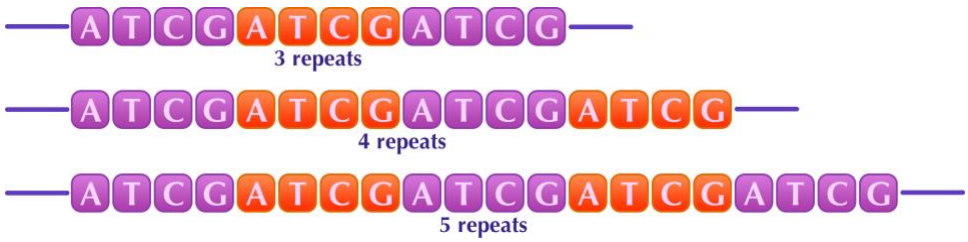

Figure 26.6. Short tandem repeats (STRs) are short, identical DNA sequences that repeat consecutively several times.

26.7.2 SHORT TANDEM REPEATS AS BIOMARKERS

The frequency of short tandem repeats (STRs) is significantly higher in the general population. These tandem repeats comprise over 50% of the human genome and notably exceed the prevalence of single nucleotide polymorphisms (SNPs) [9]. Consequently, STRs are often employed as biomarkers to monitor inheritance patterns due to their widespread presence in the general population [10].

26.7.3 MULTIPLE-LOCUS VARIABLE NUMBER TANDEM REPEATS AS BIOMARKERS

Multiple-locus VNTR fingerprinting (MLVF) employs VNTRs from various loci as biomarkers for identification. It requires calculating the repeat numbers for each locus. However, it does not allow for a straightforward and unambiguous determination of individual repeat counts at each locus. The primary drawback of this method is the absence of direct result comparisons among laboratories. Conventional electrophoresis on low-resolution agarose gels presents amplicons only as banding patterns and does not provide accurate repeat counts or correlations to PCR targets [3].

26.8 EPIDEMIOLOGICAL STUDIES CONNECTING STRS TO DISEASES

Epidemiological studies often clarify the genetic link to human diseases through disease-associated STRs. Examples include the genome-wide search for common STRs related to the genetic risk of Parkinson's disease [11] and the variants of the VNTR of the DRD4 gene associated with novelty-seeking personality traits and ADHD (attention-deficit hyperactivity disorder). The number of tandem repeats can trace the ancestral origin of these variants linked to ADHD and determine whether their prevalence results from chance or selection.

26.8.1 CASE EXAMPLE: LINKING DRD4-7R VNTRS TO ADHD

When the number of repeats is exactly seven, the 7-repeat phenotypic expression of the DRD4-7R allele is strongly associated with personality traits like novelty-seeking and ADHD (attention-deficit hyperactivity disorder) among European Caucasians, South Americans, and individuals of Middle Eastern descent [12]. In the general population, the number of tandem repeats for the DRD4 gene, which codes for dopamine D_4 receptors, ranges from 2 to 11.

26.8.1.1 The Number of DRD4 Repeats in the General Population

The number of repeats in the general population ranges from 2R to 11R for the 48-bp VNTR variants of the DRD4 gene, situated near the telomere of chromosome 11p

in exon 3, which encodes the dopamine D_4 receptor. The variants 2R, 4R, and 7R comprise over 90% of allelic diversity, with 4R being the most prevalent, whereas the frequencies of 2R and 7R vary significantly depending on geographic location [13,14].

26.8.2 THE USE OF HOMOZYGOUS VARIANTS TO TRACE ANCESTRAL HISTORIES

The frequency of occurrence can trace specific ancestral events that lead to the inheritance of identical gene variants from both parents. For the DRD4 gene, the most common homozygous variants are 2R/2R, 4R/4R, and 7R/7R, which help determine whether these specific variants arise by chance or through non-random selection. This polymorphism also assists in determining allele ages, indicating when the genetic lineage began to diverge.

26.8.2.1 Tracing the Divergence of the Ancestral 2R, 4R, and 7R Variants

In other words, one can trace the events that led to the divergence of the 2R, 4R, and 7R variants within specific subpopulations in certain geographic regions. This outcome also highlights the migration patterns in human history that have shaped the current demographic distribution. This finding is a clear example of the genealogical methodologies and analyses required to trace the ancestral events that caused the divergence of the genetic traits of the DRD4 gene variants.

26.8.2.2 Linkage Disequilibrium Analysis

Linkage disequilibrium (LD) refers to the non-random association of alleles at different loci, serving as a sensitive indicator of the population genetic forces shaping a genome [15]. The LD between alleles and traits facilitates fine-scale gene mapping through genome-wide association studies, aiding in identifying SNPs linked to complex diseases [15].

26.8.3 NATURAL SELECTION OR RANDOM OCCURRENCES

No single statistic effectively quantifies linkage disequilibrium (LD); however, local and genome-wide LD patterns provide insights into natural selection and historical population dynamics [16]. Recombination patterns can reveal complex interactions among selection, mutation, and genetic drift, all of which influence LD levels. High local LD indicates a recently favored allele under strong selection, while low LD suggests random selection without preference in mate choice. LD patterns can aid in identifying selected loci and estimating allele ages [15].

26.8.3.1 Evaluating the Selection Preferences of 2R/2R and 4R/4R

Drawing on ancestry from Africa, Europe, Asia, North and South America, and the Pacific Islands, the data for 2R/2R and 4R/4R homozygotes did not achieve statistical

significance. It showed minimal linkage disequilibrium (LD). Most other prevalent DRD4 VNTR variants do not demonstrate a selection preference, suggesting that parental mate selection is random for these 2R/2R and 4R/4R alleles.

26.8.3.2 Evaluating the Selection Preferences of 7R/7R

In contrast, the 7R/7R homozygote data revealed statistical significance through Tajima's D test [17], highlighting a strong linkage disequilibrium (LD) at most polymorphic sites. The evidence for robust LD surrounding the 7R allele is compelling, as all 7R/7R individuals (including those from Africa) exhibit a strong selection preference from both parents.

26.8.3.3 Evaluating the Recombination Pattern of 7R/7R Selection

The recombination pattern of the homogeneous 7R/7R allele suggests that selection plays a significant role in inheriting identical gene variants from both parents. The 7R variant likely originated from a rare mutation before becoming widespread due to positive selection, indicating that this selection was not a chance occurrence.

26.8.3.4 Selection of Mating Pairs for 7R Variant Traits

Evidence suggests that strong selection has raised the allele frequency to levels surpassing those expected from random genetic drift. The selection of 7R traits by mating pairs may result in higher birth and survival rates than other pairs. Beyond random selection, the greater prevalence of the 7R variant in human populations likely arises from positive selection and the survival advantages associated with this variant's gene expression.

26.8.4 THE NON-RANDOM SELECTION REFLECTED BY HOMOZYGOUS 7R/7R INHERITED FROM BOTH PARENTS

The non-random selection suggests evolutionary advantages that enhance survival rates in phenotypic traits such as novelty-seeking and ADHD associated with the 7R variant. This variant may confer benefits that have persisted despite the detrimental effects of a defective gene variant on regulating impulsive behaviors due to the D_4 receptors' insensitivity to dopamine signals.

26.8.4.1 The Survival Advantages of Non-Random Selection for DRD4-7R Traits

Dopamine's role in suppressing choices during decision-making is crucial for managing impulses against distracting stimuli. This regulation of decision-making distinguishes humans in their intellectual evolution. However, the resulting dysregulation, which leads to impulsive and novelty-seeking behaviors, paradoxically increases survival

rates compared to the general population. Novelty-seeking may have encouraged human ancestors to migrate, explore other continents, and reduce resource competition during their migrations, thereby enhancing the group's survival rate.

26.8.5 THE COMMON DISORDER-COMMON VARIANT HYPOTHESIS

The common disorder-common variant hypothesis (CDCVH) posits that if a heritable disease is prevalent, with a prevalence exceeding 1% to 5% in the population, its genetic contributors will also be widespread [18]. Genetic markers are identified in variants of coding and regulatory regions. This hypothesis applies to specific variants that increase susceptibility to complex polygenic diseases, with each variant contributing a small additive effect that combines into multiplicative effects on disease phenotypes [18].

26.8.5.1 The Paradoxical Benefits of DRD4-7R in ADHD as a Disorder

The link between the 7R allele and ADHD suggests that both environmental and genetic factors may influence this prevalent disorder. The statistical prevalence of ADHD in the general population ranges from 3% to 5%. Traits associated with DRD4-7R could predispose individuals with ADHD to behaviors that may be harmful in certain environments but beneficial in others. While the common disorder-common variant hypothesis indicates deleterious effects related to a genetic disorder, the 7R traits might promote an evolutionarily advantageous strategy that results in positive outcomes rather than the negative behavioral consequences typically associated with the common disorder.

26.8.5.2 Single-Step Mutation Random Events for Most DRD4 Variants

Polymorphic variation is observed across 67 haplotype variants. Most haplotype variants, such as 2R-6R, involve a single variant resulting from a one-step mutation. These likely represent random occurrences due to mutations or recombination errors.

26.8.6 MULTI-STEP MUTATION EVENT PATTERN FOR DRD4-7R VARIANTS

The DRD4-7R variant shows a pattern associated with nearby polymorphisms. The connection between 7R/7R homozygotes and four adjacent DRD4 polymorphisms suggests that they have likely undergone at least six mutations or recombination events. This sets it apart from other common variants that contain a single mutation.

26.8.6.1 Evidence of a More Recent Origin for Multi-Step Mutation in DRD4-7R

Although rare, this pattern likely spread through non-random mate selection for specific traits, resulting in a higher prevalence than random selection for other alleles. This prevalence prompted an examination of the linkage disequilibrium (LD) between the 4R and 7R alleles. The rare multi-step mutation in 7R/7R homozygotes probably occurred more recently than in other homozygotes, such as 4R/4R, which underwent a single one-step mutation. The recombination pattern at these polymorphic sites aligns with the selection pattern at DRD4-7R.

26.8.7 ALLELE AGE CALCULATIONS

Estimating the age of an allele offers insight into when the variant first emerged. This equation summarizes the standard methods for calculating allele age, as represented in generations:

$$t = \frac{1}{ln\,(1-c)} ln\,\frac{x(t)-y}{1-y} \tag{26-1}$$

where c is the recombination rate, $x(t)$ is the frequency in generation t, and y is the frequency on ancestral chromosomes. The age of the allele is calculated using the formula above when the recombination rate c is available, and $x(t)$ and y are derived from the genetic survey data [19].

26.8.7.1 Dating the Ages of Most Other DRD4 VNTR Variants

Calculating the allele age based on the high global prevalence of DRD4-2R, 4R, and 7R suggests that these alleles have an ancient origin, estimated to range from approximately 300,000 to 500,000 years [14]. However, the branching of the common ancestor for these variants differs in allele age.

26.8.7.2 Determining the Age of the Divergence of DRD4-7R from Its Common Ancestor

Strong linkage disequilibrium (LD) exists between the 7R allele and nearby DRD4 polymorphisms, suggesting that it is 5 to 10 times younger than the common 4R allele. Based on the variability of 18 observed intra-allelic heterozygosity sites across the locus, the age of the 7R allele is estimated to be approximately ten times younger. An analysis of high-heterozygosity sites indicates that the most recent common ancestor of the 7R/7R allele emerged between 40,000 and 50,000 years ago [13].

26.8.7.3 Likely Mate Selection Preferences of DRD4-7R Variants

The divergence from their common ancestors likely predates human migration to other continents during the Upper Paleolithic period [14]. Traits such as novelty-seeking may offer survival advantages by encouraging adventurous migration to areas with less competition for resources and mate selection. Their mate selection preferences likely reinforce themselves when both parents carry the variant traits, ensuring these traits are passed down to their descendants. With a subpopulation of either homozygous or heterozygous carriers isolated from the general population, the frequency of allele prevalence could quickly exceed chance levels due to these self-reinforcing conditions.

26.8.8 THE DRD4 GENE ENCODES THE DOPAMINE D4 RECEPTOR

The functional significance of the number of repeated sequences lies in the changes to the dopamine D_4 receptors (DRD4) protein within a region that couples with G proteins and mediates intercellular cAMP levels [20]. Consequently, the expression of the 7-repeat gene leads to the insensitivity of dopamine D_4 receptors to inhibitory signals during neural processing. This 7-repeat allele of DRD4-7R is associated with the expression of novelty-seeking personality traits and the prevalence of ADHD.

26.8.8.1 D_4 receptors in executive function control for task completion

D_4 receptors are predominantly localized in prefrontal cortex (PFC) neurons, where they play a crucial role in regulating behavior and evaluating the importance of choices during decision-making. These receptors in postsynaptic neurons respond to inhibitory dopamine signals released from presynaptic neurons by binding to them, thereby suppressing alternative options in the decision-making process. This mechanism is part of the executive functions that help control impulsive behavior, which can lead to poor choices. The executive function process involves the interaction of cognitive, behavioral, and emotional regulation for making appropriate decisions, much like a CEO's role within the brain. These receptors manage selective attention by focusing on problem-solving while maintaining working memory to keep track of essential intermediate steps for effective planning toward achieving a desired goal.

26.8.8.2 The Role of Inhibition in Decision-Making Processes

A decision involves choosing one option from several alternatives. Completing a complex task requires decision-making to identify the most crucial options and working memory to retain temporary information vital for task completion. Neurons in the PFC prioritize these options based on their importance, facilitating the selection of the top choice on the list. However, if some options are equally important, a decision must be made to resolve the tie; otherwise, one may become trapped in indecision.

26.8.8.3 Conflicts in Decision-Making

A conflict arises when options are mutually exclusive; choosing one negates the other. For instance, a conflict occurs when you want to have your cake and eat it too. Resolving a conflict requires setting aside competing options to make a desirable choice. It involves suppressing alternatives to enable the selection of the preferred option. If the other options are not properly suppressed, it will result in an impulsive decision that is incompatible with the desired solution.

26.8.8.4 The Conflict Resolution Process

Dopamine functions as the inhibitory signal that prevents the selection of undesirable options during decision-making. Presynaptic neurons in the prefrontal cortex release dopamine to suppress the responses of postsynaptic neurons, enabling them to disregard irrelevant options. When D_4 receptors become insensitive due to defects in the DRD4-7R variant, they fail to respond to the inhibitory signal, leading to poor decisions and behaviors such as distraction, impulsivity, and novelty-seeking—key symptoms of ADHD.

26.8.8.5 The Pathology of Inheriting the DRD4 7-Repeat Sequences

Indecision arises when neurons fail to resolve conflicts between choices, causing them to become stuck in indecision. Impulsivity occurs when they select an inappropriate response that seems significant as a choice despite being otherwise. Distraction happens when they focus on trivial stimuli that compete for attention. This scenario resembles driving a car with malfunctioning brakes, which can lead to crashes due to failed stopping attempts. This example illustrates the consequences of variations in the number of repeated DNA sequences that affect the sensitivity of a receptor binding site, thereby influencing behavior and neural processing. However, the prevalence statistics persist, suggesting that the disorder may confer beneficial effects on survival, in contrast to the assumed detrimental impacts of a disorder associated with novelty-seeking and curiosity behaviors, which are crucial for innovation, discovery, and migration away from resource competition — factors that likely provide a competitive advantage in survival.

26.8.8.6 The Unforeseen Impact of a Gene Variant on Global Migration and Technological Innovation

This case example illustrates that tracing ancestral lineage has revealed unexpected results from inheriting a gene variant that may confer beneficial effects despite the behavioral deficits of impulsivity and distraction. The timing of the ancestral gene variant divergence coincided with the onset of global migration, which preceded the subsequent discovery of the New World and the innovations associated with the technological advancements that began during the Upper Paleolithic.

26.9 MICROHAPLOTYPES CONSIST OF TWO OR MORE SNPS

Microhaplotypes (MHs), often called microhaps, are chromosomal segments that typically consist of two or more closely located SNPs at the molecular level (less than 300 bp) [21]. When two or more SNPs are positioned near each other, their alleles can be inherited together from parent to child. The variants often cluster together.

26.9.1 THE USE OF MICROHAPLOTYPES AS BIOMARKERS

Because of the short lengths of microhaplotypes, these sequences are often used as genetic markers. They are called "micro" because they cover only a few dozen base pairs instead of thousands. Microhaplotypes (MHs) have proven to be highly informative in demonstrating the uniqueness of DNA profiles and determining an individual's biogeographic ancestry. Biological relationship tests like paternity tests frequently use MHs as biomarkers.

26.9.1.1 The Use of SNPs, STRs, and MHs as Biomarkers

Although the polymorphism analyses for each type of biomarker are similar regarding their occurrence frequency within the target population, notable differences arise in the sequences of variations [6]. The SNP variants are limited to four types, as only four nucleotide bases can substitute: A, C, G, and T. Thus, there are four SNP variants. In contrast, the STR variants depend on the number of repeats, with probabilities of occurrence based on this count within the sequence. As the name suggests, VNTR indicates a variable number of tandem repeats, and the repeated sequence varies among individuals in the population. Identifying the number of repeats in a sample may not be precise when using amplification techniques. It is essential to recognize that false positives may arise due to sample contamination, which usually results from carryover in prior DNA analyses.

26.9.1.2 The Use of Non-Coding Regions in Forensic Analysis

Forensic analysis focuses on connecting individuals rather than solely emphasizing gene expression or the effects of protein synthesis from coding regions. Excluding coding regions from DNA sequences during forensic analysis can simplify the complexities of phenotypic expression, which may impact the results. Therefore, polymorphisms in non-coding areas are typically preferred for DNA analysis. Most DNA sequence databases used in forensic analyses primarily consist of non-coding regions.

26.10 THE FORENSIC COMBINED DNA INDEX SYSTEM DATABASES

The Combined DNA Index System (CoDIS) is a national DNA index system consisting of a collection of databases developed and maintained by the United States Federal Bureau of Investigation (FBI) to collect data on various short tandem repeats (STRs). This system distinguishes individuals based on the lengths of these alleles. The database uses a set of 20 loci tested together to improve the reliability of the analysis [22]. A locus refers to a specific location on a chromosome for a particular gene or genetic marker. The twentieth locus in CoDIS, known as AMEL, is used to identify the sex of individuals based on the sex chromosomes [6].

26.10.1 THE USE OF 20 LOCI AS INDEX FOR IDENTIFICATION

Using 20 loci significantly reduces the risk of confusing two distinct individuals, as individuals often share alleles at specific loci, especially when related. It is analogous to using physical characteristics for identification: the more traits used to describe a person, the less likely they are to be mistaken for someone else.

26.10.1.1 Forensic Index System Databases

The Forensic Index includes profiles derived from biological evidence at crime scenes, such as blood or semen samples, which help identify remains when linked to a known source [23]. DNA profiles in each index are compared to find direct matches or potential relationships based on shared genetic data. Searches for familial relationships are conducted solely for missing persons and unidentified remains.

26.10.1.2 The Unidentified Human Remains Index

The Unidentified Human Remains index features profiles of human remains. In contrast, the Relatives of Missing Persons (RMP) index includes voluntarily collected profiles from biological relatives, typically obtained from buccal swabs but sometimes sourced from blood or other samples. These profiles are indexed in this database.

26.10.1.3 The Pedigree Tree Index

The Pedigree Tree index organizes relatives' specimens into family groups for efficient searches and is compared only to the Unidentified Human Remains index in CoDIS. Samples from missing persons are stored in the Missing Person index and may include personal items such as toothbrushes or hairbrushes.

26.10.1.4 The Convicted Offender Index

The Convicted Offender Index contains profiles of individuals convicted of qualifying offenses. According to state laws, blood or buccal swabs are collected from arrestees. Some states also collect additional samples.

26.10.2 THE USE OF 13 CORE STR LOCI

A reduced number of loci is sufficient for identifying an individual without requiring all 20 loci. The FBI has identified 13 core STR loci as adequate for individual identification in the U.S. Conversely, Interpol (the International Criminal Police Organization) has determined that 10 core STR loci are sufficient for identification in the U.K. and Europe. Interpol is an intergovernmental organization composed of 196 member countries that facilitates police collaboration across these nations. In the U.S., the 13-STR profile is a standard identification method used to determine human remains, establish paternity, or link suspects to crime scenes.

26.10.2.1 Using 13 Core STR Loci for Identification

The FBI determined the frequency of each allele for the 13 core STRs across various ethnic groups by analyzing DNA from hundreds of unrelated individuals. Assuming all 13 STRs exhibit independent permutations, statistical calculations indicate that the probability of two unrelated Caucasians sharing identical STR profiles, or "DNA fingerprints," is approximately 1 in 575 trillion [24].

26.10.2.2 The Matching Fallacy

However, this probability pertains to pairs of individuals globally. With 100 million Caucasians, there are 5,000 trillion pairs, suggesting that approximately eight or nine pairs would match at the 13 STR loci. This matching does not indicate which profiles two individuals share, and the likelihood of matching a crime-related profile remains very low [25].

26.10.2.3 A STR Match

A laboratory analyzes the allele profiles of 13 core STRs from both samples to connect evidence from a crime scene to a suspect. If the STR alleles do not match, the individual is excluded as a potential source of the evidence. If they match all 13 STRs, a statistical calculation estimates how frequently this genotype occurs in the population, considering the prevalence of each STR allele in the individual's ethnic group. A Hardy-Weinberg calculation determines the frequency of the observed genotype for each STR, and multiplying the frequencies of the individual STR genotypes provides the overall profile frequency.

26.10.2.4 Analysis of Conditional Probabilities for STR Matching

If Suspect A is excluded as the source of the crime scene sample, and Suspect B matches all 13 STRs, the likelihood that a random member of Suspect B's ethnic group shares this genotype is extremely low. It implies the probability of observing this DNA profile if the evidence did not originate from the suspect. Misinterpreting this could confuse the probability of the suspect belonging to the ethnic group with the likelihood that the suspect is the source. Calculating the transitional probability requires Bayes' theorem and prior conditional probabilities concerning the suspect's involvement. Moreover, the probability increases significantly if a relative of the suspect is the source, particularly a sibling.

26.10.2.5 The Likelihood of a False Positive

The significance of chance phenomena cannot be overstated; even with an extremely low likelihood of a false positive, it can still occur, much like winning the mega lottery. Although the odds are one in a billion, winning is possible if you are lucky; if it never happens, no one could win the lottery.

26.10.2.6 The Confounding Variables

DNA collected from crime scene evidence is often limited in quantity, poorly preserved, or degraded, leading to partial profiles. Analyzing fewer than 13 STR loci increases the likelihood of random matches. For example, examining fewer than five initial STRs significantly raises the chances of encountering that genotype. Additional evidence regarding Suspect B should be gathered to exclude potential candidate involvement. Moreover, common STR alleles within an individual's ethnic group can raise genotype frequencies, even when all core loci are analyzed.

26.10.2.7 The Sources of Uncertainty

Crime scene samples may also contain DNA from multiple sources, complicating the analysis. Traditional forensic analysis typically relies on probabilities, which means that even a confirmed match cannot establish guilt. Additional evidence is necessary to establish the connection beyond a single source.

26.11 THE ARTIFICIAL INTELLIGENCE APPROACH

Artificial intelligence serves as an alternative method for solving complex problems through expert systems or machine learning capabilities. Recently, AI has gained recognition for employing machine learning to tackle intricate issues via examples rather than explicit instructional methods. The system discovers solutions by learning from the examples in the training dataset.

26.11.1 MACHINE LEARNING

AI provides alternatives to traditional methods for calculating conditional probabilities and establishing relationships. It effectively addresses specific inquiries in natural language by utilizing large language models (LLMs) to generate suitable genealogical responses to user questions. Additionally, it categorizes historical records of both written and spoken language and census records and stores genealogical data in databases. Furthermore, machine learning speeds up genome analyzers, producing millions of short sequencing reads [26]. However, it is essential to recognize AI's strengths and weaknesses when evaluating the validity of conclusions.

26.11.1.1 Automation with Machine Intelligence

Today's AI is more accurately described as machine intelligence (MI) because it forms internal representations through learning instead of relying solely on preprogrammed responses. Rather than adhering to strict algorithms, these models transform internal representations from input to output using neural networks that extract information from training datasets.

26.11.1.2 Collective Neural Processing in Computing

The system generates responses using a vast network of neurons, unlike the single CPU logic found in traditional computers. AI output improves through machine learning, beginning with random outputs refined through extensive training on billions of examples to establish essential input-output relationships. Forming meaningful connections among data points in a dataset relies on a highly interconnected network of neurons that captures relationships by creating an internal representation of these links.

26.11.1.3 How an AI system is trained

An AI neural network generates outputs through interconnected tokens and connection weights. It responds to queries based on these tokens and weight matrices derived from extensive training datasets, effectively mapping inputs to outputs. Learning from billions of examples creates an internal representation of input-output relationships. The network applies learning rules to modify interconnectivity and uses criteria to identify relevant connections while discarding irrelevant ones.

26.11.1.4 The Mystery and Magic of AI Uncovered

Requesting an AI to generate images from text involves mapping inputs to outputs through training to minimize errors. This approach aids the system in adjusting its input-output connections. Pairing descriptions with images form internal representations that enhance learning by reducing errors. Trial and error produce random outputs, which are compared to target images. A reduction in errors indicates progress, prompting the

system to modify its connections accordingly. Repeating this process with a billion examples refines the image based on the description and reinforces relevant connections. The outcome is an image created from the text description.

26.11.1.5 AI Responses to User Queries

A genealogical query can yield plausible results when trained on questions with known answers, utilizing billions of examples to develop appropriate responses. Training an AI system generally requires billions of examples for the model's representation to align with expected input-output relationships. One unique feature of modern AI systems is their ability to generalize and summarize large sets of data or information.

26.11.2 HOW AI PRODUCES RESPONSES TO UNFAMILIAR QUESTIONS

It can generate appropriate responses even if those inquiries are not included in the training dataset. It generalizes answers using internal interconnectivity to map input queries to output response relationships. This means it can provide missing information, even when the database is incomplete or the inquiries are not part of the original training examples.

26.11.2.1 Consequences of Insufficient Training

Adequate training on a large dataset can produce reliable responses. However, these responses may become inaccurate if the internal model fails to converge properly to a stable representation due to insufficient training data. An LLM model typically requires a minimum dataset size of 200 billion examples to achieve convergence toward a solution.

If the training dataset is insufficient, AI models may struggle to converge and become trapped in local minima during error minimization. Techniques are available to help escape these local minima and find better global solutions, although errors may occur while overcoming the hump. Without these methods, AI tools might inadvertently settle for suboptimal solutions.

26.11.3 REQUIREMENTS FOR DNA DATASET IN AI TRAINING

AI for DNA analysis relies on short polymorphisms or genome-wide sequences stored in databases. Training an AI system using human genome data composed of 3 billion nucleotides is possible. Training on short polymorphic DNA sequences requires a substantial database to provide sufficient example datasets. GenBank is an open-access, annotated repository of nucleotide sequences and their corresponding protein translations. It includes 34 trillion base pairs from over 4.7 billion sequences and covers over 580,000 protein species. It contains 1.42 million SNPs, averaging one every 1.9 kb.

Approximately 60,000 SNPs are present in exons (both coding and untranslated regions), with 85% located within 5 kb of the nearest SNP [9].

26.11.3.1 Exploring Relationships in DNA Datasets

AI genealogy and forensic analyses depend on accessible datasets. When searching for candidate genes associated with diseases or drug responses, reasonable suggestions are accepted even without thorough justifications. Researchers investigate potential interactions of gene expressions using extensive SNP datasets or genomes, refining their search for additional human analysis. Such tasks are labor-intensive and time-consuming. AI can accelerate the discovery of possible gene combinations, as the matching criteria are less stringent than those used for tracing an individual.

26.11.3.2 Collective Neural Processing in Computing

AI generates human-like responses by analyzing vast amounts of data; however, it falls short in logical reasoning and genuine creativity. Although this process can be automated, it lacks true intelligence, reasoning skills, and a fundamental understanding of the world. It produces content based on training datasets, which can lead to absurd outputs resulting from misrepresentations of physical phenomena.

26.12 THE ARTIFICIAL INTELLIGENCE COMPUTING APPROACHES

AI has evolved into two approaches to automated computing, each based on distinct methodologies for problem-solving. The term “AI” originated as a theoretical concept for information processing in the 1950s, a function that was once exclusive to humans before the emergence of automated computing machines. A historical case study will demonstrate how AI has advanced from the 1950s to modern-day AI.

26.12.1 CASE HISTORY: THE TURING MACHINE

Alan Turing was a brilliant mathematician recognized as a pioneer in laying the foundation for modern computing. In 1936, while at the National Physical Laboratory in the UK, he published a design for the ACE (Automatic Computing Engine) [27,28], an early version of modern computers. He proposed using mechanized computation to automate a series of programmable instructions known as the “Universal Turing Machine” [29].

26.12.1.1 The Predecessors of Modern Computing Machines

In perspective, the slide rule was a mechanical device used to calculate parabolic trajectories while accounting for recoil on naval ships by solving the differential equations of its time. Mechanical devices were employed to encode and decode wartime messages during WWII. The Colossus computers, modeled after the ACE, operated from 1943 to

1945 and were used for the cryptanalysis of the Lorenz cipher. As the first programmable computers, they utilized vacuum tubes to process Boolean logic operations. The renowned EDVAC, designed by John von Neumann in 1945, was based on Turing's theoretical work [30].

26.12.1.2 The Turing Test for Artificial Intelligence

AI dates back to 1950 when Alan Turing introduced the "imitation game" as a test for machine intelligence in his paper "Computing Machinery and Intelligence" [31]. The test, known as the "Turing Test," poses the question, "Can a machine think?" If a machine's responses are indistinguishable from those of a human, it is considered intelligent. However, this test evaluates performance rather than cognitive ability. For example, a person with schizophrenia might be mistaken for a machine due to irrational responses, indicating that mimicking human interaction does not equate to true intelligence. Similarly, a zombie can imitate interactions but lacks conscious intelligence.

26.12.1.3 Machine Intelligence as an Automated System

The Turing Test for thinking machines has sparked philosophical discussions that extend beyond intelligence to encompass consciousness. Ultimately, the criteria for imitation are insufficient to differentiate consciously thinking machines from those that are unintelligent. Intelligence requires thinking, while imitation does not. The Turing Test fails to address this crucial aspect, leading to ongoing debates about intelligence that are both unnecessary and unproductive — debates that persist even today.

26.12.2 MACHINE LEARNING WITHOUT PREPROGRAMMED ALGORITHMS

Machine intelligence refers to automated systems that use machine learning (ML) to process information without predefined algorithms, achieving performance levels similar to human interaction. ML emphasizes computational processes rather than outcomes. Artificial intelligence encompasses abilities that can surpass human capabilities.

26.12.2.1 Decoding Messages: A Prelude to AI

Turing worked at Bletchley Park in London during WWII, focusing on cracking the Enigma machine code used by German forces [32]. Initially, Polish mathematicians decrypted Enigma messages; however, the Germans changed the cipher daily. He and others at Bletchley conducted cryptanalysis using the Bombe machine and the Banburismus technique to decode messages. They successfully deciphered encrypted naval communications from German U-boats and the Air Force in 1941. In July 1942, he developed the Turingery technique to decrypt the top-secret German communications utilized in the Lorenz cipher machine.

26.12.2.2 The Demand for Artificial Intelligence in Encryption

He traveled to the U.S. in December 1942 to share his expertise in Enigma encryption and advise U.S. military intelligence. This decryption was crucial in directing Allied convoys away from the U-boat wolf packs, significantly changing the Battle of the Atlantic in favor of the Allies during WWII [32].

26.12.2.3 Post-War Recognition of Decryption Technology

During his visit, Turing learned about the U.S. speech encoding system and developed his speech-scrambling device, "Delilah." In 1945, he was awarded the Most Excellent Order of the British Empire (OBE) for his contributions during the war. Established by King George V in 1917, the OBE honors achievements in the arts, sciences, charitable work, and public service in non-combat roles [33].

26.12.2.4 The Official Apology

In 1952, Alan Turing was prosecuted for homosexuality, which was illegal in Britain at the time. To avoid imprisonment, he accepted chemical castration and had his security clearance revoked. In 1954, he died from cyanide poisoning at the age of 41, a death ruled as suicide. Turing's role in breaking the Enigma code remained classified until the 1970s, with the complete story emerging in the 1990s. His conviction was overturned posthumously in 2013, 60 years later, and he received a Royal Prerogative of Mercy pardon from the Queen.

26.12.2.5 Turing Award for Excellence in Computing

Turing's contributions to computer science are recognized through the annual Turing Award, which has been the highest honor in the field since 1966 and is regarded as equivalent to a Nobel Prize. Notably, the Nobel Foundation did not establish a category for computer science awards in 1895 [34].

26.13 TWO DISTINCT AI APPROACHES FOR SOLVING COMPLEX PROBLEMS

As AI has developed through advanced computing algorithms capable of solving complex problems, two fundamentally different approaches have emerged. The traditional approach employs preprogrammed algorithms to outline specific methodologies for addressing expert system problems. Conversely, the alternative approach utilizes machine learning to uncover solutions to complex issues without specifying the precise methods needed to solve them.

26.13.1 TRADITIONAL ALGORITHM METHOD

This approach utilizes symbolic processing and algorithms to automate tasks in expert systems. Traditional DNA analysis employs algorithmic methods within expert systems to calculate conditional probabilities through database comparisons. This illustrates the classical approach to problem-solving by distinctly outlining the methodologies for computer automation.

26.13.1.1 Special Purpose Machines

This methodology addresses problems by following the algorithm's explicit instructions. The intelligence resides with the programmer, not the machine itself. Expert systems are designed to address well-defined issues in specific scenarios within a targeted domain. They are not meant to serve as universal problem solvers for other issues and do not generalize solutions to different problems.

26.13.1.2 Limitations Imposed by Its Preprogrammed Algorithms

Using programmed logic to tackle complex problems may seem straightforward and effective, but these algorithms limit their problem-solving capabilities. Unforeseen circumstances that surpass the programmed logic remain unaddressed unless the algorithms are modified. They cannot incorporate new algorithms or analyses without reprogramming.

26.13.1.3 Unresolvable Missing or Incomplete Information

Another limitation is its inability to address information gaps when the database is incomplete, missing, or contains biased data. The programmer serves as the primary source of intelligence rather than the machine. The machine executes automated instructions without the intellectual capacity to generate new algorithms or independently rectify errors.

26.13.1.4 The Unresolved "Hard Problems" in Natural Language and Image Processing

For decades, computers have struggled to recognize speech accents in automated phone answering systems and transcribe voicemails. Facial recognition, handwriting recognition, and natural language processing exemplify the limitations of traditional AI. These shortcomings persist because traditional AI relies on pre-set algorithms that cannot comprehend anything beyond their programming. Furthermore, chatbots, which do not need to process accents, cannot interpret context or extract relevant information from users 'inquiries to provide appropriate responses.

26.13.2 NEURAL NETWORK MACHINE LEARNING TECHNIQUES

The neural network approach uses implicit methods to derive solutions from examples without depending on explicit algorithms. The Perceptron, an early problem-solving machine, classifies images by learning from training data. Modern AI systems utilize similar learning rules and neural network architectures from the 1980s, scaled and trained on datasets of 100 billion, enabling them to address real-world challenges rather than merely solving toy problems.

26.13.2.1 Problem-Solving Using Neural Networks

Recently, machine learning techniques that utilize neural networks have greatly enhanced natural language processing and image recognition. These methods enable systems to accurately route calls to the appropriate department for human responses and identify faces, handwriting, and speech accents through extensive training on datasets.

26.13.2.2 General Purpose Problem Solvers

Machine learning acquires and generalizes solutions to various problems, establishing itself as a universal problem solver. It can also address gaps created by missing data not included in the training set. This system is trainable and adapts to the training datasets. With additional training, it can refine its responses; thus, it is an adaptive system that gradually aligns with the user's environment.

26.13.3 PRINCIPLES OF NEURAL INTEGRATION: FROM MULTIPLE INPUTS

Neural networks adjust their weights based on training feedback and compute outputs across multiple neurons. They employ parallel processing to generate solutions from a group of neurons, relying on statistical properties instead of the accuracy of each neuron. This contrasts with the serial processing found in traditional CPUs, which utilize algorithmic problem-solving.

26.13.3.1 Learning Rules for Adjusting Connection Weights

Learning rules decrease errors or employ the auto-associative Hebbian learning rule, reinforcing active connections to improve network performance. The learning rule used during training can be either a supervised approach (learning with a teacher) or an unsupervised approach (learning without a teacher). Incremental weight adjustments enable the system to find solutions that minimize errors; therefore, these solutions arise automatically through error minimization.

26.13.3.2 Limitations of Single-Layer Neural Networks

The original Perceptron network architecture was overly simplistic, relying on a single-layer design to solve problems. This limitation caused researchers in the traditional algorithmic AI group to underestimate its problem-solving abilities, claiming that it could not tackle real-world challenges. However, this limitation was soon recognized as an inherent flaw of a single-layer neural network. In contrast, a multi-layer network architecture can tackle problems that a single-layer network cannot solve.

26.13.3.3 Overcoming Limitations Using Multi-Layer Neural Networks

Subsequently, it was found that adopting a multi-layer network architecture was crucial for addressing these complex issues. Multi-layer network models effectively overcame the limitations of single-layer neural networks. They tackled problems previously deemed "hard" by the expert system algorithms of that time, such as backing up a tractor-trailer. Nevertheless, limited computing resources led to their dismissal for many years as mere solutions to "toy problems," even though they demonstrated the capacity to solve unforeseen challenges with limited computing resources.

26.13.3.4 Recent Popularization of Neural Computing

The renewed interest in AI and machine learning has recently demonstrated its practicality by training on a vast dataset compiled from over a billion websites. The multi-layered network architecture and learning principles in deep learning are fundamentally the same as those used decades ago, only scaled up a billionfold. They utilize the same supervised training and reinforcement learning models developed over fifty years ago, applied to a training dataset that is billions of times larger.

26.13.4 COMPUTATIONAL RESOURCES NEEDED FOR NEURAL COMPUTING

This technology has become a powerful tool for delivering solutions that rival human capabilities. For instance, ChatGPT employs large language models (LLMs) to identify patterns in text and generate human-like conversations. It was trained on data from a billion websites, utilizing 500 billion tokens, 175 billion parameters, and extensive weight matrices to create a complex neural network. This framework encompasses nearly a trillion variables, crucial for computation and storage, facilitating the establishment relationships among higher-order text strings as tokens.

The computing resources are vast, encompassing speed, memory, and storage. Generating text dialogues comparable to human conversations required a month of calculations on cloud server farms that were previously accessible only to major tech companies. Although these dialogues appear related to the topic, they may lack relevance without proper context. Current AI models excel at summarizing large datasets by

generalizing based on the higher-order correlation of input-output relationships; however, they do not yet possess the capability to perform logical reasoning or deduction.

26.13.4.1 Generating Credible Responses from Patterns without Reasoning

With the computing resources available in cloud server banks, neural networks can summarize large volumes of complex data, such as analyzing combinations of genetic profiles from extensive DNA datasets to identify interrelated relationships. However, these connections represent plausible relationships that may not be based on logic but rather on patterns. Therefore, AI-generated genetic matches are plausible but not necessarily valid.

26.13.4.2 Legal Criteria for Validity

Legal criteria require that forensic analysis exceeds reasonable doubt. Concerns regarding AI-generated relationships arise because they may appear valid based on random associations instead of logical reasoning. Therefore, they do not inherently satisfy the legal standards for conviction or determining parenthood matches.

26.13.4.3 Less Restrictive Requirements for Plausible Heritage

Exploring heritage can be an enjoyable exercise in genealogical tracing. It seeks to uncover potential lineage rather than adhere to strict criteria for definitive results. Therefore, using AI to suggest possible heritage is only a recommendation, even if the lineage is hypothetical rather than guaranteed. Examining potential heritage can be a rewarding experience.

26.13.5 BRIDGING THE INFORMATION GAPS

Unlike traditional expert systems, AI trained through machine learning can identify connections even when there are gaps in the training data. This capability effectively fills information voids by suggesting plausible links rather than leaving them unexplored as unknown.

26.13.5.1 Unexplained Conclusions

AI machine learning systems connect inputs to outputs by identifying relationships within training datasets. By examining the input-output patterns observed during training, we can draw conclusions. However, this method often lacks transparency, making understanding how outputs are generated from billions of examples without a logical explanation is difficult.

The connection may be plausible, but it lacks logic. Without understanding the reasoning behind the conclusions, precautions should verify the validity of the results.

However, even without a strong rationale, the results can still support exploratory analysis and highlight specific features for human assessment without entirely dismissing them.

26.13.6 EXPLAINABLE AI (XAI)

Explainable AI (XAI) research clarifies the reasoning behind AI responses by aligning response patterns with logical explanations. This approach promotes interpretable conclusions. It enhances confidence in their validity rather than leaving them unexplainable.

26.14 CASE HISTORY: ARTIFICIAL NEURAL NETWORKS

A historical overview of artificial neural networks and artificial neurons in computing is essential for understanding AI systems that exhibit human-like intelligence. Understanding these principles enhances our trust in AI-generated outcomes and deepens our comprehension of genetics and inheritance. Trainable artificial neural systems depend on established physical and mathematical computational principles rather than mystical ones.

26.14.1 BIO-INSPIRED NEURON COMPUTATION

The theoretical foundation for neural processing in neural networks is based on biological neurons' anatomical and physiological principles of biological neurons. Artificial neurons were inspired by the works of neuroanatomist Santiago Ramón y Cajal, neurophysiologist Sir Charles Scott Sherrington, and neuropsychologist William James. Their theories linked the fundamental functions of neurons to behaviors, establishing a connection between the mind and body.

26.14.1.1 Case Example: The McCulloch-Pitts Neuron

In 1943, Warren McCulloch and Walter Pitts [35] demonstrated that neurons can perform logical operations through threshold activation functions. A perceptron is a binary classifier that uses supervised learning rules to derive solutions to classification problems from training data without explicit instructions.

26.14.1.2 Case Example: The Perceptron Neuron

Artificial neurons (see Figure **26.7**) have shown the ability to perform the logical OR operation with a threshold of 1 (see Table 1) and the logical AND operation with a threshold of 2 (see Table 2). In this model, inputs and outputs are represented as 1 (true) or 0 (false), consistent with the original Perceptron model.

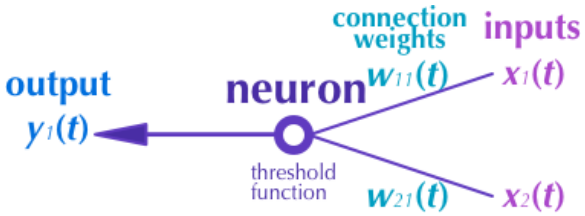

Figure 26.7. A diagram illustrating a McCulloch-Pitts neuron highlights the relationships between inputs, outputs, and connection weights. The connection weight calculates the weighted sum of the neuron's inputs. The threshold activation function determines the neuron's output, activating when the weighted sum reaches a specific threshold. It computes the logical OR function when the threshold is set to 1 and the logical AND function when set to 2, with both inputs and outputs represented as 1 (true) or 0 (false).

26.14.1.3 Computing the Boolean Logical OR Operation with Artificial Neurons

A truth table (see Table 1) illustrates the connection weights that define input-output relationships. In this example, the connection weights are set to 1. By adjusting these weights, machines can learn from examples without explicit instructions on solving problems. When the threshold is set to 1, the neuron computes the Boolean logical OR operation.

Logical OR Operation Using a Threshold of 1						
Input, $x_1(t)$	**Input, $x_1(t)$**	**Weight, $w_{11}(t)$**	**Weight, $w_{11}(t)$**	**Weighted Sum, $\Sigma = w_{11}(t) + w_{21}(t)$**	**Threshold, $\theta = 1$**	**Output, $y_1(t)$**
0	**0**	1	1	**0**	$\Sigma < \theta, y_1(t) = \mathbf{0}$	**0**
0	**1**	1	1	**1**	$\Sigma \geq \theta, y_1(t) = \mathbf{1}$	**1**
0	**0**	1	1	**1**	$\Sigma \geq \theta, y_1(t) = \mathbf{1}$	**1**
1	**1**	1	1	**2**	$\Sigma \geq \theta, y_1(t) = \mathbf{1}$	**1**

Table 1. This truth table illustrates the neural computation of the logical OR operation with a threshold of 1. The output activates when the weighted sum of inputs exceeds this threshold, with inputs and outputs represented as 1 (true) or 0 (false).

26.14.1.4 Computing the Boolean Logical AND Operation with Artificial Neurons

When the threshold is set to 2, the truth table (see Table 2) illustrates that it computes the Boolean logical AND operation. This concept was instrumental in using artificial neurons for AI computation. They can also compute other logical NOT (negation), NOR (NOT OR), and NAND (NOT AND) operations.

Logical AND Operation Using a Threshold of 2						
Input, $x_1(t)$	**Input, $x_1(t)$**	**Weight, $w_{11}(t)$**	**Weight, $w_{11}(t)$**	**Weighted Sum, $\Sigma = w_{11}(t) + w_{21}(t)$**	**Threshold, $\theta = 2$**	**Output, $y_1(t)$**
0	**0**	1	1	**0**	$\Sigma < \theta, y_1(t) = 0$	**0**
0	**1**	1	1	**1**	$\Sigma < \theta, y_1(t) = 0$	**0**
0	**0**	1	1	**1**	$\Sigma < \theta, y_1(t) = 0$	**0**
1	**1**	1	1	**2**	$\Sigma \geq \theta, y_1(t) = 1$	**1**

Table 2. This truth table illustrates the neural computation of the logical AND operation with a threshold of 2.

26.14.1.5 Case Example: The Perceptron Machine

In 1958, Frank Rosenblatt developed a Perceptron machine for image processing, enabling it to classify images through training rather than programmed instructions [36]. It could distinguish between cats and non-cats without needing explicit directions. Although it initially generated excitement in AI research, the capabilities of neural computation faced criticism due to its single-layer architecture, particularly when it could not solve problems like XOR computation.

26.14.1.6 The AI Winter

In 1969, Marvin Minsky and Seymour Papert published a book [37]critiquing the limitations of perceptrons because of their inability to solve the XOR computation. As a result, funding for neural network research dwindled for decades. The algorithmic approach was preferred over machine learning for addressing complex problems.

26.14.2 THE DOMINANCE OF TRADITIONAL ALGORITHMIC EXPERT SYSTEMS

A common issue is the lack of sufficient DNA data in databases that compute matches, such as the missing genetic profile information in a genetic index database. Traditional algorithmic AI approaches have been the gold standard for expert systems in addressing these challenges for decades, even though they often struggle with straightforward problems humans can easily solve. These “hard” problems remain unresolved because no practical, explicit algorithmic solutions exist.

26.14.2.1 The ‘‘Hard” Problems in Algorithmic AI

Examples of challenging problems for traditional AI include filling in missing information in expert systems, such as addressing gaps in DNA databases to trace ancestry until a neural network is used to supplement the missing profile information in those databases. Other challenging problems that seem obvious for humans to solve include speaker-dependent speech recognition, scale- and rotation-invariant handwriting recognition, and orientation- and viewpoint-invariant facial recognition in image analysis.

These complex issues are now solvable through machine learning. Neural computation identifies patterns in training datasets to generalize solutions without needing algorithmic instructions.

26.14.3 COMPUTATIONS USING ARTIFICIAL NEURONS

An artificial neuron processes input signals from multiple sources to produce a single output (see Figure **26.7**). It is analogous to a biological neuron, which receives thousands of synaptic inputs through its dendrites and transmits output via a single axonal pathway to other neurons. The application of artificial neurons for computation was inspired by the anatomical structure of biological neurons, which possess dendritic branches resembling tree branches.

26.14.3.1 Simultaneous Parallel Processing with Artificial Neurons

The design of neural computing uses neurons to gather inputs from thousands of other neurons, performing weighted-sum operations concurrently rather than sequentially. Each neuron operates as an independent processor, allowing it to process information simultaneously with other neurons in the network. This emphasizes the distinction between parallel processing in neural computation and serial processing in traditional digital CPUs.

26.14.3.2 Differences Between Neural and Conventional Digital Processing

Neural computing processes thousands of inputs simultaneously and in parallel, while conventional digital computing handles multiple inputs sequentially, addressing one or two at a time. This serial method creates bottlenecks at the CPU when dealing with a large number of inputs. The speed advantage in neural computing arises from processing signals in parallel instead of in series.

26.14.3.3 The Neural Processing Chip Hardware Accelerator

Neural computing can be computationally expensive for large-scale matrix multiplications involving floating-point and logical operations (see sections below). Conventional digital computing alleviates CPU bottlenecks by enhancing processing speed. However, modern AI systems utilizing large-scale LLM models analyze billions of inputs; for instance, processing 100 billion mathematical operations with a trillion parameters at 10 GHz takes 10 seconds for a single step in this context. Training a general-purpose neural network requires billions of iterative steps before the system converges on a potential solution. Estimates indicate that power consumption by AI servers will exceed 20% of total power consumption in the U.S. within the next five years. Nevertheless, a smaller, special-purpose AI system designed for customized solutions will consume significantly fewer resources and necessitate less training data.

26.14.3.4 Computational Bottlenecks

To alleviate bottlenecks, neural computing chips are designed to execute matrix operations as a unified computing unit instead of sequentially on a conventional CPU. Furthermore, processing thousands of inputs necessitates a substantial memory footprint, which requires on-chip memory banks. Analog computing, similar to biological neurons, could significantly enhance speed and reduce power consumption.

26.14.3.5 Accelerating Computational Speed with Analog Signals in Neurons

Biological neurons evolved to overcome computational bottlenecks by employing analog signal processing. They process signals electrically through membrane potentials, integrating thousands of inputs as analog signals. Excitatory inputs generate positive potentials, while inhibitory inputs create negative potentials for summation. The membrane potentials aggregate synaptic inputs as graded potentials.

26.14.4 MATHEMATICAL COMPUTATION IN BIOLOGICAL NEURONS

They perform mathematical operations for addition and subtraction using both positive and negative potentials. The attenuation adjusts the weighted sum by scaling the potential amplitude along the path to the integration zone. Signals are attenuated based on the impedance of the signal path and the distance from the synapse to the axon hillock. Most signal paths from a synapse follow a dendritic branch unless the synapses are located on the soma.

26.14.4.1 Nonlinear Operations as a Threshold in Biological Neurons

The firing threshold defines the threshold function for activation based on whether the weighted sum surpasses it. The gating mechanisms of the ionic channels embedded in the membrane convert analog signals into digital signals in the form of action potentials, which are transmitted to the next neuron for processing.

26.14.4.2 Digital Sigal Transmission in Biological Neurons

To maintain signal transmission integrity, neurons transmit digital pulse-coded signals encoded by action potentials over long distances along the axonal membrane. Upon reaching the axon terminal, these action potential signals are converted into chemical signals that act as neurotransmitters. This system primarily functions as a hybrid for processing electrochemical signals. By analogy, most batteries also operate as hybrid electrochemical systems.

26.14.4.3 Chemical Sgnal Transmission in Biological Neurons

Before transmitting signals to the next neuron, neurons convert electrical signals into chemical signals in the form of neurotransmitter packets that encode excitatory or inhibitory responses. An excitatory response increases the likelihood of the next neuron's firing, while an inhibitory response decreases this probability. This neurotransmitter release process acts as a universal mechanism for encoding both types of signals to regulate the firing of the subsequent neuron.

26.14.4.4 Transmission of excitatory and inhibitory signals in neurons

An excitatory synapse uses specific neurotransmitters as chemical signals that bind to receptors, converting these pulse codes into a positive membrane potential in the next neuron and facilitating addition. Conversely, an inhibitory synapse employs a different neurotransmitter that binds to a distinct receptor, transforming the signals into a negative potential in the subsequent neuron and enabling subtraction.

26.14.4.5 Parallel Processing in Biological Neural Networks

Each neuron processes both analog and digital signals using the same mechanisms, accelerating the overall process by concurrently managing signals among billions of neurons in the brain. This capability enables massive parallel neural computing, effectively addressing the limitations of serial computing. Although the processing speed of biological neurons is a million times slower than that of electronic computers, the brain can simultaneously compute at high speeds by processing computations across billions of neurons.

26.14.5 THRESHOLD PROCESSING IN BIOLOGICAL NEURONS

The gating mechanisms of ionic channels at the trigger zone initiate action potentials at the axon hillock. When the summed membrane potential exceeds this threshold, the weighted-sum analog signal is converted into a digital signal as an action potential.

26.14.5.1 Digital Signal Processing in Biological Neurons

The action potential is a digital pulse signal produced by a neuron. It is characterized by a consistent amplitude and width, which encode the neural response as a time series of action potentials. The neuron transmits these digital signals along the axon as outputs for the next neuron to process.

26.14.5.2 Hybrid Chemical and Electrical Signal Processing in Biological Neurons

The signal is converted into a chemical signal through neurotransmitters released at the synapse. This chemical signal can be either excitatory or inhibitory, depending on whether it transmits a go or no-go signal to continue or block communication. The processing advances to the next layers of neurons, resulting in collective rather than individual computation by neurons.

26.14.5.3 Voltage-Gated Channels as Differential Equation Solvers for Biological Neurons

The firing threshold is regulated by a set of ionic channels embedded in the membrane. Each ionic channel in a neuron has voltage gates that govern a series of differential equations related to voltage activation. A typical neuron contains sodium (Na) channels and potassium (K) channels that regulate this threshold [38]. As Alan Hodgkin and Andrew Huxley described, these thresholds were controlled by gating channels and rectifiers between 1948 and 1952 [39–45]. Each neuron effectively processes a set of differential equations through analog signal processing, driven by the regulation of threshold voltage. Hodgkin and Huxley were awarded the Nobel Prize in 1963 for their contributions to neural computation, particularly for discovering the gating properties of these ionic channels [46] before their existence was confirmed.

26.14.5.4 Neural Computation with Voltage-Gated Channels

Erwin Neher and Bert Sakmann later employed the patch-clamp technique to confirm the existence of gating currents generated by these ionic channels. They were also awarded the Nobel Prize in 1991 [47]. This exemplifies a case of discovering biological neural computations that inspired processing with artificial neurons.

26.14.5.5 Synaptic Relays in Biological Neurons

The synapse is where electrical signals travel between neurons across a narrow gap. Neurotransmitters transport these signals across the gap, converting them into voltage in the following neuron. Biological neurons process information using electrical signals, relaying results to the next neuron through a complex transformation into chemical signals before returning to electrical signals for further processing.

Each connection influences firing, with certain connections having a greater impact due to their synaptic strength and proximity to the soma (cell body). A neuron activates when the total synaptic inputs at the axon hillock exceed a threshold, generating action potentials (nerve impulses).

26.14.6 MATHEMATICAL WEIGHTED SUM IN BIOLOGICAL NEURONS

A neuron acts as an adder, summing inputs from earlier layers in the neural network. When this computational process occurs for all neurons in a network, signals are processed progressively along the pathways of each neuron in parallel. In other words, as conventional computers do, neural computation distributes processing across all neurons without depending on a single CPU.

26.14.7 NEURONS AS A VOTING SYSTEM: COUNTING VOTES FROM SYNAPTIC INPUTS

Each connection to a neuron carries a weight that influences its responses. When all weights are equal, the neuron receives unbiased inputs. It operates like a voting system, gathering input votes to determine the output based on a threshold. This threshold signals whether the votes are sufficient to declare a winner, similar to a simple majority or a supermajority in decision-making. The neuron activates its output when the total count exceeds the threshold, reflecting the outcome of the voting process.

26.14.7.1 The Use of Input Bias in Neural Computation

In a fair election, each vote is counted equally. This means all connection weights are unbiased and treated the same when calculating the total. However, if some connection weights are greater than others, those votes hold more significance. Biological neurons manipulate biases in vote counting to adjust the importance of specific inputs for learning and survival.

26.14.7.2 Adjusting Relevant Inputs via Connection Weights

To survive, one must focus on essential stimuli and ignore irrelevant ones. If all stimuli were equally significant, sensory overload could overwhelm an individual. Therefore, learning emphasizes inputs related to survival while filtering out unimportant ones. By adjusting synaptic weights based on learning principles, the system hones in on crucial stimuli and dismisses those that provide less relevant information.

26.14.8 THE “USE IT OR LOSE IT” PRINCIPLE OF SYNAPTIC PLASTICITY

Biological neurons utilize Hebbian learning to modify synaptic weights. This activity-dependent rule reinforces connections when neural inputs are activated simultaneously. It reflects the “use it or lose it” principle. Cognitive decline in Alzheimer’s patients is mitigated when they participate in mental activities.

26.14.8.1 The Auto-Associative Hebbian Learning Rule

The Hebbian learning rule is an auto-associative learning mechanism employed for unsupervised learning, meaning it does not require a "teacher" to train the system. It learns to associate relevant inputs that are activated simultaneously. The closer the temporal proximity, the stronger the association.

26.14.8.2 Learning Rules for Modifying Connection Weights

The learning rule is simple: strengthen important connections and diminish less significant ones. By modifying connection weights, the system prioritizes responses to critical stimuli. Over time, this process slowly enhances suitable responses by emphasizing positive feedback while ignoring negative feedback. As a result, adjusting connection weights to favor selected inputs improves our overall effectiveness.

26.14.8.3 The Choice of Feedback for Adjusting Connection Weights

Determining which stimuli to prioritize depends on the value of the feedback. Similarly, neurons select specific types of feedback to guide the learning rules for adjusting their connection weights. Because some feedback is more effective than others, this decision ultimately influences the results of appropriate responses.

26.14.9 PRINCIPLES OF NEURAL COMPUTATION: THE WEIGHTED SUM

If $x_i(t)$ represents the i-th input to the neuron at time t, and $w_i(t)$ denotes the connection weights (see Figure **26.7** and Figure 26.9), then the weighted sum $s(t)$ is computed as follows:

$$s(t) = w_1(t)x_1(t) + \cdots + w_i(t)x_i(t) \cdots + w_n(t)x_n(t) = \sum_{i=1}^{i=n} w_i(t)x_i(t) \tag{26-2}$$

26.14.9.1 The Vector Representation of Neural Inputs

This indicates that an input vector $\vec{\boldsymbol{X}}(\boldsymbol{t})$ represents all n inputs to the neuron:

$$\vec{\boldsymbol{X}}(\boldsymbol{t}) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \tag{26-3}$$

26.14.10PRINCIPLES OF MATRIX MULTIPLICATION IN A NETWORK LAYER

In a simple layer of m input neurons, each with n inputs and $w_{ij}(t)$ representing the connection weights of the j-th neuron, the weighted sum is computed using the following matrix equation:

$$\begin{bmatrix} s_1(t) \\ \vdots \\ s_m(t) \end{bmatrix} = \begin{bmatrix} w_{11}(t) & \cdots & w_{n1}(t) \\ \vdots & \ddots & \vdots \\ w_{1m}(t) & \cdots & w_{nm}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \tag{26-4}$$

26.14.11PRINCIPLES OF INTERNAL REPRESENTATION BY WEIGHT MATRICES

Represent the above equation using matrix notation as follows:

$$\vec{\boldsymbol{S}}(\boldsymbol{t}) = \vec{\boldsymbol{W}}(\boldsymbol{t}) \cdot \vec{\boldsymbol{X}}(\boldsymbol{t}) \tag{26-5}$$

The equation above illustrates the matrix multiplication performed by each neuron. The matrix $\vec{\boldsymbol{W}}(\boldsymbol{t})$ represents the connection weights $w_{ij}(t)$ from the i-th neuron to the j-th neuron during the matrix multiplication process:

$$\vec{\boldsymbol{W}}(\boldsymbol{t}) = \begin{bmatrix} w_{11}(t) & \cdots & w_{n1}(t) \\ \vdots & \ddots & \vdots \\ w_{1m}(t) & \cdots & w_{nm}(t) \end{bmatrix} \tag{26-6}$$

This matrix computation occurs before a threshold function is applied to determine the output. If the weighted sum exceeds this threshold, the neuron activates its output and transmits activation information to the next neuron for further processing. If it does not exceed the threshold, the neuron remains inactive until the next time step, repeating the same process.

26.14.12THE SIZE PRINCIPLE OF NEURAL NETWORKS

A biological neuron typically receives about 10,000 to 100,000 synaptic inputs from other neurons and performs weighted-sum calculations in a single step that takes less than a millisecond. The human brain contains approximately 100 billion neurons, each with an average of 10,000 to 100,000 inputs, and its computational requirements are comparable to those of current AI systems.

26.14.12.1 The One-Step Processing Principle in Neural Computation

Processing speeds improve when all neurons carry out these mathematical operations simultaneously instead of one node at a time. In today's AI networks, each artificial neuron can have over a trillion parameters to update at each time step. The bottleneck arises from the serial processing performed on conventional multi-core CPUs.

26.14.12.2 Bio-inspired Neuron-Based Weighted Sum Computation

Biological neurons inspire neural processing in artificial neurons. Each biological neuron has thousands of synaptic inputs on its dendritic branches and soma, resembling the leaves of a tree branch. It integrates synaptic potentials at the axon hillock to determine whether to fire, but only if the summed voltage reaches a certain threshold. It enhances matrix computation by summing voltages as analog signals from distant synaptic inputs, weighted by the attenuation along the dendritic tree branch.

26.14.12.3 Computing Uniform Contributions from Inputs

If all the weights are equal, the system simplifies to a straightforward sum of the inputs. Each input contributes equally, resulting in an unbiased system. Neural computation becomes a simple "adder" operation, similar to how a computer performs addition within the CPU. The computational processes in both digital and neural computing are analogous, illustrating the similarities between these two mechanized computations.

26.14.12.4 Principles for Adjusting Connection Weights According to Input Contributions

If each connection weight changes individually, the neuron adjusts the contribution of each input according to its updated weight. It gradually adapts to the significance of each input's contribution, leading to learning through the reorganization of its connections with other neurons.

26.14.13 PRINCIPLES OF COMPUTATION FOR CORRELATION FUNCTIONS

Higher connection weights significantly enhance the weighted sum, resulting in a positive correlation. Conversely, negative weights diminish the sum, leading to a negative correlation. Time-delayed temporal processing computes the correlation function for a time series of pulse-coded action potentials [48]. The temporal correlation functions generated by biological neurons emerge from their processing of spike trains, which creates the association between input signals represented by an internal correlation matrix [49–52]. Neurons perform temporal integration of synaptic inputs for processing [53,54].

This process enables neurons to learn to generate appropriate outputs by recognizing higher-order correlations for each connection, even if these are indirectly related.

26.14.13.1 Principles of Hebbian Learning Rule

As explained earlier, the learning rule gradually adjusts weights, allowing the system to respond to relevant inputs for the desired output. The adjustment rule is straightforward: strengthen the connection by increasing the weight when the input is relevant and decrease it when it is not. This process modifies the influence of relevant inputs on neuron firing when activated. The connection is weakened for irrelevant inputs by lowering the weight when not activated.

26.14.13.2 Principles of Emphasizing Contributions of Relevant Inputs by Weights

The neuron adjusts its connection weights over time, amplifying relevant inputs and diminishing irrelevant ones. Eventually, its weights stabilize to reflect the importance of these connections. Only relevant inputs affect firing based on their weighted sums, while irrelevant inputs have minimal influence.

26.14.14PRINCIPLE OF THE OUTPUT ACTIVATION HARD THRESHOLD STEP FUNCTION

The threshold activation function acts as a step function for an artificial neuron (see Figure 26.8). It is discrete and discontinuous, defined by a precise threshold. At this threshold, differentiation becomes undefined. As a nonlinear function, it cannot be traced back from the output to the original inputs. The output activation changes abruptly as a step function. It is 1 when the weighted sum exceeds the threshold; otherwise, it is 0:

$$y_j(t) = \begin{cases} 1, \text{if} \sum_{i=1}^{i=n} w_{ij}(t)x_i(t) \geq \theta \\ 0, \text{if} \sum_{i=1}^{i=n} w_{ij}(t)x_i(t) < \theta \end{cases} \quad (26\text{-}7)$$

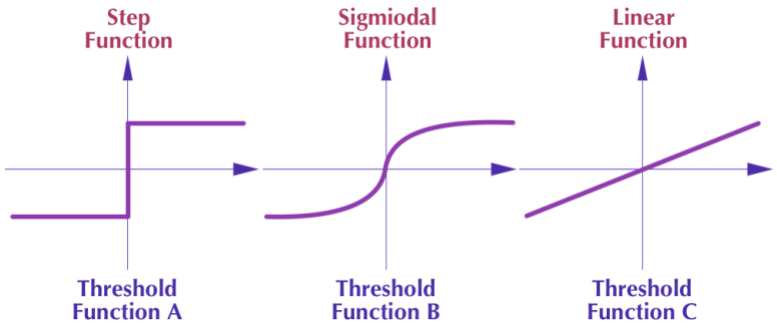

Figure 26.8. Graphs illustrating various types of threshold functions. A step function threshold is a discrete, discontinuous function that indicates a hard threshold. A sigmoidal function threshold is a continuous function that signifies a soft threshold. A linear function threshold effectively represents the absence of a threshold.

26.14.14.1 The Sigmoid Soft Threshold Activation Function

Other threshold functions, such as the sigmoidal function (see Figure 26.8), provide a gradual change in output activation rather than a discrete step change. The threshold activation function is smooth and continuous, acting as a soft threshold. It is a nonlinear function that cannot be reversed from the output to the original inputs.

26.14.14.2 Principles for Collapsing Linear Threshold Functions into a Single-Layer Network

A linear threshold (see Figure 26.8) represents a continuous function equivalent to no threshold. As a linear function, it does not change the output activation produced by the weighted sum. The activation function equals the weighted sum of the inputs. In such cases, the series of matrix multiplications can be simplified into a single equivalent matrix of a single-layer network, which does not perform any significant processing.

26.14.15PRINCIPLES OF MULTI-LAYER NEURAL NETWORKS

A multi-layer neural network consists of hidden layers situated between the input and output layers (see Figure 26.9). Each hidden layer processes and extracts essential information to address more complex tasks. Each non-linear layer abstracts information related to the input-output relationships, facilitating effective feature extraction. As the network learns, it gradually adjusts its internal representation of these relationships. The learning rule modifies the connection weights according to the significance of these input-output relationships. Since these relationships are expressed as matrices, they encapsulate the internal model representation. Each hidden layer captures a unique representation, extracting different connections to provide higher-order input-output relationships.

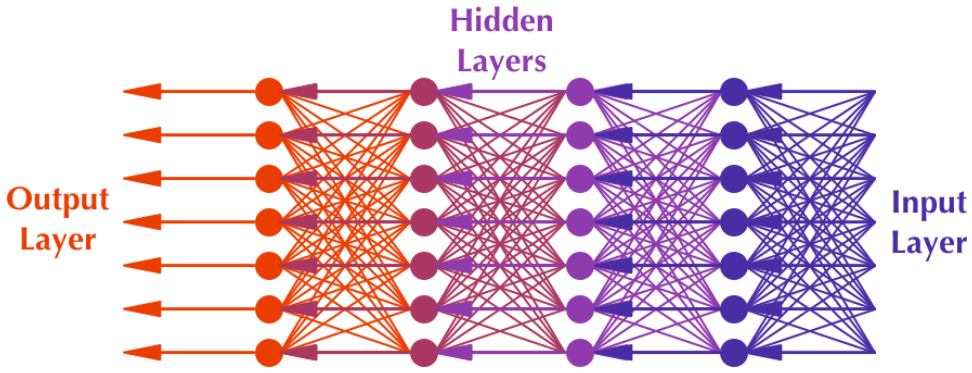

Figure 26.9. A multi-layer neural network illustrating the different layers and their input-output relationships. The hidden layers between the input and output layers capture complex connections among the inputs and outputs.

26.14.15.1 Principles of Matrix Multiplication in Neural Computation

Representing the weighted sum of inputs with a matrix results in an output activation as follows:

$$y_j(t) = \begin{cases} 1, \text{if} \begin{bmatrix} w_{11}(t) & \cdots & w_{n1}(t) \\ \vdots & \ddots & \vdots \\ w_{1m}(t) & \cdots & w_{nm}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \geq \theta \\ 0, \text{if} \begin{bmatrix} w_{11}(t) & \cdots & w_{n1}(t) \\ \vdots & \ddots & \vdots \\ w_{1m}(t) & \cdots & w_{nm}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} < \theta \end{cases} \tag{26-8}$$

26.14.15.2 Principles of Computed Output Activation Functions

By utilizing matrix notation to simplify the matrices, the output activation function can be expressed by the following equation:

$$\vec{\boldsymbol{Y}} = \begin{cases} 1, & \text{if} \vec{\boldsymbol{W}} \cdot \vec{\boldsymbol{X}} \geq \theta \\ 0, & \text{if} \vec{\boldsymbol{W}} \cdot \vec{\boldsymbol{X}} < \theta \end{cases} \tag{26-9}$$

26.14.15.3 Principles of Computation with Irreplaceable Hidden Layers

By cascading the outputs of one layer to the next, the neurons in the intermediate layers can use the output of each neuron as input. The intermediate layer plays a crucial role in neural computation and signal processing. It addresses the limitations of models like the perceptron, which cannot solve the XOR problem due to its single-layer structure. Hidden layers facilitate non-linear processing through the threshold function; without them, computations may revert to a single layer.

26.14.15.4 Principles of Nonlinear Neural Network Computation

The threshold function, a binary step function that performs Boolean logic, is mathematically nonlinear. This nonlinearity prevents the series of matrix multiplications from simplifying into a single equivalent multiplication. If it were otherwise, the computations of a multilayer network would be reducible to those of a single-layer neural network's matrix multiplication.

26.14.15.5 The Essential Processing by Intermediate Layers

The hidden layers are essential for computations in neural networks. Intermediate interneurons would be unnecessary if a single-layer equivalent could replace the network. Due to nonlinear functions, these layers compute critical behaviors that simpler networks cannot replicate.

26.14.15.6 The Essential Processing of Multilayer Neural Networks

Logical inference and causal reasoning in network connections stem from nonlinear neural processing. These processes are irreversible due to the noncommutativity of matrix multiplication, which renders it impossible to trace neural networks back to the origins of consciousness's abilities. This distinction distinguishes AI from traditional preprogrammed algorithmic approaches to autonomously solving intellectual processing functions.

26.14.15.7 Principles of Methodologies for Adjusting Internal Representations

The principle involves learning rules to adjust connection weights based on network feedback to minimize errors. These five processes will impact the network's performance:

- The quality and quantity of the dataset for training a network;
- The learning rules for updating the connection weights between neurons;
- The selection of relevant feedback for determining the above updates;
- The network criteria for improving the outcomes of the network outputs;
- The methodology for avoiding convergence into a non-optimal solution.

26.14.16 PRINCIPLES OF THE OPTIMIZATION PROCESS IN AI TRAINING

Machine learning in AI models is a mathematical, data-driven methodology that optimizes solutions by minimizing errors to align with expected results. It employs a trial-and-error approach to learn from mistakes. By reducing discrepancies between the

system's output and the anticipated results in each iteration, the system typically converges on a solution that achieves the desired outcomes.

26.14.16.1 Principles of the Generalization Process in AI

Generalizing a solution requires considerable time and numerous training examples to determine expected outcomes by establishing extensive relevant connections. This process fosters an internal representation that enables generalization from a training dataset, even if the input query is not part of that dataset. The key lies in the internal representation created by a network of neural connections that establishes higher-order correlations in the input-output relationships.

26.14.17EVALUATION OF AI NETWORK PERFORMANCE

Understanding AI model training is crucial for selecting a model for genealogical analysis. A network's performance hinges on the learning rules used during training and its connectivity to determine relevance. This comprehension enables an objective evaluation rather than relying solely on vendor marketing. Below is a summary of AI methodologies for assessing performance.

26.14.17.1 Principles of Minimizing Error Methods

The backprop rule reduces errors by progressively propagating them from the output layer to the hidden layers. The error is calculated by comparing the network's outputs with the desired outputs and is then used to update the connection weights for each neuron. This method minimizes errors through gradient descent and was developed by Cauchy in 1847 [55].

26.14.17.2 Principles of Backprop Networks Commonly Used in AI

Backprop, short for "back-propagating error correction," was coined by David Rumelhart, Geoffrey Hinton, and Ronald Williams [56,57]. Numerous predecessors contributed to its application of backprop in neural networks. Shunichi Amari suggested training multilayer perceptrons (MLPs) in 1967 [58] with end-to-end connections using stochastic gradient descent (SGD) [59]. The concept of backpropagation was introduced by Rosenblatt in 1962. The backprop model gained popularity from the book "Parallel Distributed Processing."

26.14.17.3 Limitations of the Backprop Machine Learning Model

The backprop model faces various computational challenges. A supervised learning model requires a "teacher" to provide correct answers for training. The model adapts to the training set as guided by the trainer. Training a supervised learning AI model with human input on correct responses is labor-intensive; for instance, it involves

teaching the system to differentiate between male and female faces in facial recognition. In contrast, other intelligent systems utilize self-learning through the unsupervised learning paradigm to categorize independently without any instructions via auto-association.

It uses a gradient descent algorithm to find an optimal solution but may get stuck in a local minimum (see below). Further methods are needed for the system to ascend before descending to a better solution.

26.14.18THE GRADIENT DESCENT METHOD FOR ERROR MINIMIZATION

One method for solving the error minimization problem is gradient descent, which follows the gradient downhill (see Figure 26.10). However, the network may converge to a solution and become trapped in a local minimum, representing a non-optimal solution. Figure 26.10Figure 26.10 illustrates the topology of the solution space.

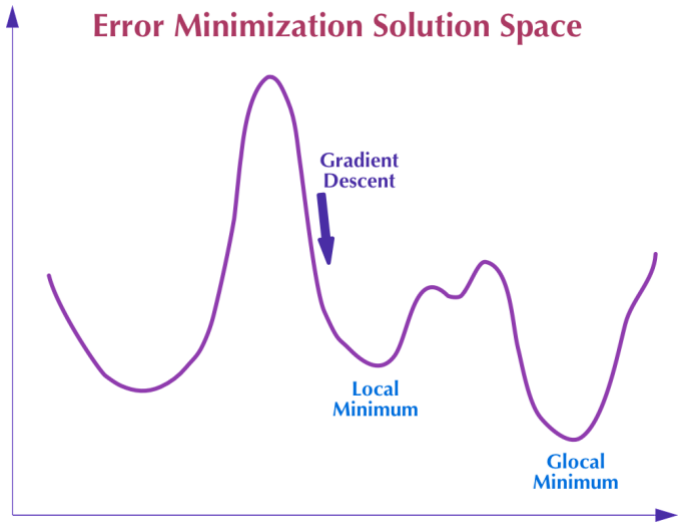

Figure 26.10. A graph illustrating gradient descent for error minimization highlights the risk of encountering local minima. Escaping these non-optimal solutions often requires counterintuitive approaches, such as moving against the gradient uphill before converging on a more optimal solution. It is often impossible to determine the location of the global minimum or to confirm whether the best solution has been found without prior knowledge of that optimal solution.

The discovered solution may be valid but might not be optimal. Often, one must ascend before descending into another minimum, which requires using alternative methods to counter the gradient descent algorithm. However, achieving a global minimum is not guaranteed if the optimal solution is unknown beforehand.

26.14.18.1 Methods for Escaping Local Minima

Various analytical methods aim to escape local minima, often using stochastic and probabilistic output activation functions. Semi-random fluctuations help ascend against the gradient toward an optimal solution. This approach generates random outputs to explore nearby solution spaces while moving against the downward gradient. It employs stochastic techniques rather than deterministic ones. Stochastic processing is probabilistic, solving problems differently from conventional deterministic approaches that specify exact solutions without variations. Some stochastic solutions even tunnel through the hill to escape local minima.

26.14.19CASE EXAMPLE: THE BOLTZMANN MACHINE

Hinton employs the Boltzmann machine, drawing on the analogy from statistical physics regarding energy to define the exploration function in a search through the solution space. Random thermal fluctuations enable the system to acquire enough energy to escape from local minima. The temperature of the system's energy dictates the extent of exploration. Instead of the supervised learning model of backprop, the Boltzmann machine uses unsupervised learning for auto-association. Hinton received the Nobel Prize in 2024 for utilizing statistical physics energy functions to optimize solutions in the Boltzmann machine [60].

26.14.20PRINCIPLES OF SUPERVISED LEARNING TRAINING METHOD

A drawback of backprop is its dependency on supervised learning for training. This approach necessitates prior knowledge of the correct answers, which is unlike unsupervised learning, where such knowledge is not required. Moreover, training involves direct interaction and demands human guidance.

Human input is crucial to minimizing AI errors when training self-driving cars to recognize pedestrians and traffic lights. This process is labor-intensive and time-consuming. While they may seem automated, supervised learning AI models necessitate significant human training before arriving at an effective solution.

26.14.21THE UNSUPERVISED LEARNING TRAINING METHOD

The unsupervised learning model discovers solutions by forming associations between input-output pairs. Donald Hebb initially proposed the biological mechanisms for synaptic plasticity, suggesting how neurons modify synaptic strengths. The Hebbian learning rule states that simultaneous activation strengthens connections while inactivity weakens them. This process allows neurons to adjust weights without an external "teacher." The learning rule essentially establishes correlation functions autonomously, reflecting the "use it or lose it" principle introduced earlier.

26.15 SUMMARY

Genealogical and forensic analyses of genetic data are illustrated with case examples to introduce the core principles of probabilistic inheritance and analytical techniques for determining ancestral lineages and identity matches. Traditional algorithmic approaches depend on expert systems that utilize defined statistical inference methodologies to address these issues analytically; however, they cannot resolve unexpected scenarios and manage missing genetic information.

Alternative AI methodologies utilize machine learning to demonstrate their benefits and practical limitations. Key machine learning principles are summarized to provide an informed evaluation of AI tools' performance without suggesting they possess magical abilities in solving genealogical challenges. The AI approach requires extensive training on large datasets to identify suitable solutions and learns from examples to summarize data through input-output correlation. Consequently, it can bridge gaps in genealogical databases, although these outcomes may not always be logically sound.

Until AI improves its reasoning capabilities, caution is advised when drawing conclusions. However, these systems effectively automate the cataloging of genealogical databases from various records and can generate automated responses to inquiries using natural language processing to answer questions from everyday users. Understanding the mechanics of AI machine learning will assist in a thorough evaluation of the performance of different AI models for specific genealogical applications.

References

(1) Bush, W. S. Genome-Wide Association Studies. In *Encyclopedia of Bioinformatics and Computational Biology*; Ranganathan, S., Gribskov, M., Nakai, K., Schönbach, C., Eds.; Academic Press: Oxford, 2019; pp 235–241. https://doi.org/10.1016/B978-0-12-809633-8.20232-X.

(2) Ala-Korpela, M.; Kangas, A. J.; Inouye, M. Genome-Wide Association Studies and Systems Biology: Together at Last. *Trends in Genetics* **2011**, *27* (12), 493–498. https://doi.org/10.1016/j.tig.2011.09.002.

(3) Figueroa, C. J.; Tang, Y.-W.; Taur, Y. Principles and Applications of Genomic Diagnostic Techniques. In *Molecular Medical Microbiology*; Elsevier, 2015; pp 381–397. https://doi.org/10.1016/B978-0-12-397169-2.00022-6.

(4) Mansur, Y. A.; Rojano, E.; Ranea, J. A. G.; Perkins, J. R. Chapter 7 - Analyzing the Effects of Genetic Variation in Noncoding Genomic Regions. In *Precision Medicine*; Deigner, H.-P., Kohl, M., Eds.; Academic Press, 2018; pp 119–144. https://doi.org/10.1016/B978-0-12-805364-5.00007-X.

(5) Schifferdecker, K.; Richey, N. Artificial Intelligence and Genealogy: Promise and Problems. *Kentucky Libraries* **2024**, *88* (3). https://openurl.ebsco.com/EPDB%3Agcd%3A14%3A32190126/detailv2?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Agcd%3A180828763.

(6) Novroski, N. M.; Cihlar, J. C. Evolution of Single-nucleotide Polymorphism Use in Forensic Genetics. *Wiley Interdisciplinary Reviews: Forensic Science* **2022**, *4* (6), e1459. https://doi.org/10.1002/wfs2.1459.

(7) Bergström, A.; Stanton, D. W. G.; Taron, U. H.; Frantz, L.; Sinding, M.-H. S.; Ersmark, E.; Pfrengle, S.; Cassatt-Johnstone, M.; Lebrasseur, O.; Girdland-Flink, L.; Fernandes, D. M.; Ollivier, M.; Speidel, L.; Gopalakrishnan, S.; Westbury, M. V.; Ramos-Madrigal, J.; Feuerborn, T. R.; Reiter, E.; Gretzinger, J.; Münzel, S. C.; Swali, P.; Conard, N. J.; Carøe, C.; Haile, J.; Linderholm, A.; Androsov, S.; Barnes, I.; Baumann, C.; Benecke, N.; Bocherens, H.; Brace, S.; Carden, R. F.; Drucker, D. G.; Fedorov, S.; Gasparik, M.; Germonpré, M.; Grigoriev, S.; Groves, P.; Hertwig, S. T.; Ivanova, V. V.; Janssens, L.; Jennings, R. P.; Kasparov, A. K.; Kirillova, I. V.; Kurmaniyazov, I.; Kuzmin, Y. V.; Kosintsev, P. A.; Lázničková-Galetová, M.; Leduc, C.; Nikolskiy, P.; Nussbaumer, M.; O'Drisceoil, C.; Orlando, L.; Outram, A.; Pavlova, E. Y.; Perri, A. R.; Pilot, M.; Pitulko, V. V.; Plotnikov, V. V.; Protopopov, A. V.; Rehazek, A.; Sablin, M.; Seguin-Orlando, A.; Storå, J.; Verjux, C.; Zaibert, V. F.; Zazula, G.; Crombé, P.; Hansen, A. J.; Willerslev, E.; Leonard, J. A.; Götherström, A.; Pinhasi, R.; Schuenemann, V. J.; Hofreiter, M.; Gilbert, M. T. P.; Shapiro, B.; Larson, G.; Krause, J.; Dalén, L.; Skoglund, P. Grey Wolf Genomic History Reveals a Dual Ancestry of Dogs. *Nature* **2022**, *607* (7918), 313–320. https://doi.org/10.1038/s41586-022-04824-9.

(8) Gao, S.; Li, B.; Mao, L.; Wang, W.; Zou, D.; Zheng, J.; Zhou, M.; Yu, S.; Zheng, F.; Yin, Y.; Liu, S. Q.; Yang, H.; Wang, H. A Theoretical Base for Non-Invasive Prenatal Paternity Testing. *Forensic Sci Int* **2023**, *346*, 111649. https://doi.org/10.1016/j.forsciint.2023.111649.

(9) Sachidanandam, R.; Weissman, D.; Schmidt, S. C.; Kakol, J. M.; Stein, L. D.; Marth, G.; Sherry, S.; Mullikin, J. C.; Mortimore, B. J.; Willey, D. L.; Hunt, S. E.; Cole, C. G.; Coggill, P. C.; Rice, C. M.; Ning, Z.; Rogers, J.; Bentley, D. R.; Kwok, P. Y.; Mardis, E. R.; Yeh, R. T.; Schultz, B.; Cook, L.; Davenport, R.; Dante, M.; Fulton, L.; Hillier, L.; Waterston, R. H.; McPherson, J. D.; Gilman, B.; Schaffner, S.; Van Etten, W. J.; Reich, D.; Higgins, J.; Daly, M. J.; Blumenstiel, B.; Baldwin, J.; Stange-Thomann, N.; Zody, M. C.; Linton, L.; Lander, E. S.; Altshuler, D. A Map of Human Genome Sequence Variation Containing 1.42 Million Single Nucleotide Polymorphisms. *Nature* **2001**, *409* (6822), 928–933. https://doi.org/10.1038/35057149.

(10) Fan, H.; Chu, J.-Y. A Brief Review of Short Tandem Repeat Mutation. *Genomics Proteomics Bioinformatics* **2007**, *5* (1), 7–14. https://doi.org/10.1016/S1672-0229(07)60009-6.

(11) Bustos, B. I.; Billingsley, K.; Blauwendraat, C.; Gibbs, J. R.; Gan-Or, Z.; Krainc, D.; Singleton, A. B.; Lubbe, S. J. Genome-Wide Contribution of Common Short-Tandem Repeats to Parkinson's Disease Genetic Risk. *Brain* **2023**, *146* (1), 65–74. https://doi.org/10.1093/brain/awac301.

(12) Nikolaidis, A.; Gray, J. R. ADHD and the DRD4 Exon III 7-Repeat Polymorphism: An International Meta-Analysis. *Soc Cogn Affect Neurosci* **2010**, *5* (2–3), 188–193. https://doi.org/10.1093/scan/nsp049.

(13) Wang, E.; Ding, Y.-C.; Flodman, P.; Kidd, J. R.; Kidd, K. K.; Grady, D. L.; Ryder, O. A.; Spence, M. A.; Swanson, J. M.; Moyzis, R. K. The Genetic Architecture of Selection at the Human Dopamine Receptor D4 (DRD4) Gene Locus. *Am J Hum Genet* **2004**, *74* (5), 931–944. https://doi.org/10.1086/420854.

(14) Ding, Y.-C.; Chi, H.-C.; Grady, D. L.; Morishima, A.; Kidd, J. R.; Kidd, K. K.; Flodman, P.; Spence, M. A.; Schuck, S.; Swanson, J. M.; Zhang, Y.-P.; Moyzis, R. K. Evidence of Positive Selection Acting at the Human Dopamine Receptor D4 Gene Locus. *Proc Natl Acad Sci U S A* **2002**, *99* (1), 309–314. https://doi.org/10.1073/pnas.012464099.

(15) Slatkin, M. Linkage Disequilibrium--Understanding the Evolutionary Past and Mapping the Medical Future. *Nat Rev Genet* **2008**, *9* (6), 477–485. https://doi.org/10.1038/nrg2361.

(16) Kreitman, M. Methods to Detect Selection in Populations with Applications to the Human. *Annu Rev Genomics Hum Genet* **2000**, *1*, 539–559. https://doi.org/10.1146/annurev.genom.1.1.539.

(17) Tajima, F. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* **1989**, *123* (3), 585–595. https://doi.org/10.1093/genetics/123.3.585.

(18) El-Fishawy, P. Common Disease-Common Variant Hypothesis. In *Encyclopedia of Autism Spectrum Disorders*; Volkmar, F. R., Ed.; Springer New York: New York, NY, 2013; pp 719–720. https://doi.org/10.1007/978-1-4419-1698-3_1998.

(19) Slatkin, M.; Rannala, B. Estimating Allele Age. *Annu Rev Genomics Hum Genet* **2000**, *1*, 225–249. https://doi.org/10.1146/annurev.genom.1.1.225.

(20) Bonvicini, C.; Faraone, S. V.; Scassellati, C. Attention-Deficit Hyperactivity Disorder in Adults: A Systematic Review and Meta-Analysis of Genetic, Pharmacogenetic and Biochemical Studies. *Mol Psychiatry* **2016**, *21* (7), 872–884. https://doi.org/10.1038/mp.2016.74.

(21) Kidd, K. K.; Podini, D. A Brief Introduction to Microhaplotypes and Their Uses. *Medical Research Archives* **2024**, *12* (3). https://doi.org/10.18103/mra.v12i3.5142.

(22) Norrgard, K. Forensics, DNA Fingerprinting, and CODIS. *Nature Education* **2008**, *1* (1), 35. https://acces.ens-lyon.fr/acces/thematiques/evolution/accompagnement-pedagogique/accompagnement-au-lycee/premiere-2019/transmission-variation-et-expression-du-patrimoine-genetique/l2019histoire-humaine-lue-dans-son-genome/empreinte-genetique/forensics-dna

(23) Osborn-Gustavson, A. E.; McMahon, T.; Josserand, M.; Spamer, B. J. The Utilization of Databases for the Identification of Human Remains. In *New perspectives in forensic human skeletal identification*; Elsevier, 2018; pp 129–139. https://doi.org/10.1016/B978-0-12-805429-1.00012-0.

(24) Reilly, P. Legal and Public Policy Issues in DNA Forensics. *Nat Rev Genet* **2001**, *2* (4), 313–317. https://doi.org/10.1038/35066091.

(25) Weir, B. S. The Rarity of DNA Profiles. *The annals of applied statistics* **2007**, *1* (2), 358. https://doi.org/10.1214/07-AOAS128.

(26) Kircher, M.; Stenzel, U.; Kelso, J. Improved Base Calling for the Illumina Genome Analyzer Using Machine Learning Strategies. *Genome Biol* **2009**, *10* (8), R83. https://doi.org/10.1186/gb-2009-10-8-r83.

(27) Turing, A. M. On Computable Numbers, with an Application to the Entscheidungsproblem. A Correction. *Proceedings of the London Mathematical Society* **1938**, *2* (1), 544–546. https://doi.org/10.1112/plms/s2-43.6.544.

(28) Turing, A. M. On Computable Numbers, with an Application to the Entscheidungsproblem. *J. of Math* **1936**, *58* (345–363), 5. https://doi.org/10.1112/plms/s2-42.1.230.

(29) Turing, A. Computing Engine (1947). *The Essential Turing* **2004**, 362. https://www.google.com/books/edition/The_Essential_Turing/dSUTDAAAQBAJ?hl=en&gbpv=1.

(30) von Neumann, J. First Draft of a Report on the EDVAC. *IEEE Annals of the History of Computing* **1993**, *15* (4), 27–75. https://doi.org/10.1109/85.238389.

(31) Turing, A. M. I.—Computing Machinery and Intelligence. *Mind* **1950**, *LIX* (236), 433–460. https://doi.org/10.1093/mind/LIX.236.433.

(32) Imperial War Museums. *How Alan Turing Cracked the Enigma Code*. Imperial War Museums. https://www.iwm.org.uk/history/how-alan-turing-cracked-the-enigma-code.

(33) *Orders and Medals*. UK Honours System. https://honours.cabinetoffice.gov.uk/about/orders-and-medals/.

(34) *The Nobel Foundation*. NobelPrize.org. https://www.nobelprize.org/the-nobel-prize-organisation/the-nobel-foundation/.

(35) McCulloch, W. S.; Pitts, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *The bulletin of mathematical biophysics* **1943**, *5*, 115–133. https://link.springer.com/article/10.1007/BF02478259.

(36) Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review* **1958**, *65* (6), 386–408. https://doi.org/10.1037/h0042519.

(37) Minsky, M.; Papert, S. An Introduction to Computational Geometry. *Cambridge tiass., HIT* **1969**, *479* (480), 104. https://leon.bottou.org/publications/pdf/perceptrons-2017.pdf.

(38) Catterall, W. A.; Raman, I. M.; Robinson, H. P. C.; Sejnowski, T. J.; Paulsen, O. The Hodgkin-Huxley Heritage: From Channels to Circuits. *J Neurosci* **2012**, *32* (41), 14064–14073. https://doi.org/10.1523/JNEUROSCI.3403-12.2012.

(39) Hodgkin, A. L.; Huxley, A. F.; Katz, B. Measurement of Current-Voltage Relations in the Membrane of the Giant Axon of Loligo. *J Physiol* **1952**, *116* (4), 424–448. https://doi.org/10.1113/jphysiol.1952.sp004716.

(40) Armstrong, C. M.; Hollingworth, S. Na(+) and K(+) Channels: History and Structure. *Biophys J* **2021**, *120* (5), 756–763. https://doi.org/10.1016/j.bpj.2021.01.013.

(41) Hodgkin, A. L. The Local Electric Changes Associated with Repetitive Action in a Non-Medullated Axon. *J Physiol* **1948**, *107* (2), 165–181. https://doi.org/10.1113/jphysiol.1948.sp004260.

(42) Hodgkin, A. L.; Huxley, A. F. A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve. *J Physiol* **1952**, *117* (4), 500–544. https://doi.org/10.1113/jphysiol.1952.sp004764.

(43) Hodgkin, A. L.; Huxley, A. F. Currents Carried by Sodium and Potassium Ions through the Membrane of the Giant Axon of Loligo. *J Physiol* **1952**, *116* (4), 449–472. https://doi.org/10.1113/jphysiol.1952.sp004717.

(44) Hodgkin, A. L.; Huxley, A. F. The Components of Membrane Conductance in the Giant Axon of Loligo. *J Physiol* **1952**, *116* (4), 473–496. https://doi.org/10.1113/jphysiol.1952.sp004718.

(45) Hodgkin, A. L.; Huxley, A. F. The Dual Effect of Membrane Potential on Sodium Conductance in the Giant Axon of Loligo. *J Physiol* **1952**, *116* (4), 497–506. https://doi.org/10.1113/jphysiol.1952.sp004719.

(46) *The Nobel Prize in Physiology or Medicine 1963*. NobelPrize.org. https://www.nobelprize.org/prizes/medicine/1963/summary/.

(47) *The Nobel Prize in Physiology or Medicine 1991*. NobelPrize.org. https://www.nobelprize.org/prizes/medicine/1991/summary/.

(48) Tam, D. C. Interspike Interval Decoding Neural Network. Patent No. US-5216752-A, June 1, 1993. https://image-ppubs.uspto.gov/dirsearch-public/print/downloadPdf/5216752.

(49) Tam, D. C.; Perkel, D. H. A Model for Temporal Correlation of Biological Neuronal Spike Trains; 1989. https://doi.org/10.1109/IJCNN.1989.118667.

(50) Tam, D. C. Decoding of Firing Intervals in a Temporal-Coded Spike Train Using a Topographically Mapped Neural Network. In *IJCNN. International Joint Conference on Neural Networks*; 1990. https://doi.org/10.1109/IJCNN.1990.137965.

(51) Tam, D. C. A Cross-Interval Spike Train Analysis: The Correlation between Spike Generation and Temporal Integration of Doublets. *Biological Cybernetics* **1998**, *78*, 95–106. https://doi.org/10.1007/s004220050417.

(52) Fitzurka, M. A.; Tam, D. C. A Joint Interspike Interval Difference Stochastic Spike Train Analysis: Detecting Local Trends in the Temporal Firing Patterns of Single Neurons. *Biological Cybernetics* **1999**, *80* (5), 309–326. https://doi.org/10.1007/s004220050528.

(53) Tam, D. C.; Perkel, D. H.; Tucker, W. S. Temporal Correction of Multiple Neuronal Spike Trains Using the Back-Propagation Error Correction Algorithm. *Neural Networks* **1988**. https://doi.org/10.1016/0893-6080(88)90312-7.

(54) Tam, D. C. A Spike Train Analysis for Detecting Temporal Integration in Neurons. *Neurocomputing* **1999**. https://doi.org/10.1016/S0925-2312(99)00104-6.

(55) Lemaréchal, C. Cauchy and the Gradient Method. *Doc Math Extra* **2012**, *251* (254), 10. https://content.ems.press/assets/public/full-texts/books/251/chapters/online-pdf/978-3-98547-540-7-chapter-4938.pdf.

(56) Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323* (6088), 533–536. https://www.nature.com/articles/323533a0.

(57) Rumelhart, D. E.; McClelland, J. L.; Group, P. R. *Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations*; The MIT Press, 1986. https://doi.org/10.7551/mitpress/5236.001.0001.

(58) Amari, S. A Theory of Adaptive Pattern Classifiers. *IEEE Transactions on Electronic Computers* **2006**, No. 3, 299–307. https://doi.org/10.1109/PGEC.1967.264666.

(59) Robbins, H.; Monro, S. A Stochastic Approximation Method. *The annals of mathematical statistics* **1951**, 400–407. https://www.jstor.org/stable/2236626.

(60) *The Nobel Prize in Physics 2024*. NobelPrize.org. https://www.nobelprize.org/prizes/physics/2024/hinton/facts/.