

Abstract:

The genome of the COVID-19 disease was sequenced for the first time in January 2020, in Wuhan, China's capital. COVID-19 genome sequencing is important for understanding the virus's origin and mutation rate, as well as the creation of vaccines and successful prevention strategies.

The use of artificial intelligence techniques to learn details from COVID-19 genome sequences is investigated in this paper. SPM is first used on a computer-understandable corpus of COVID-19 genome sequences to see whether any hidden patterns can be discovered, such as nucleotide base patterns and their relationships.

Second, the corpus is subjected to sequence prediction models to see whether nucleotide base(s) can be predicted from previous ones. Third, an algorithm is designed to identify the positions in the genome sequences where the nucleotide bases are modified and determine the mutation rate for mutation analysis in genome sequences.

Introduction:

The SARS-CoV-2 virus was discovered in a pneumonia patient in Wuhan, China's capital.

According to the World Health Organization's most recent survey, the COVID-19 has affected more than 65 million people, with almost half of them dying. 1.5 million people have died as a result of the epidemic, which has now spread to over 200 countries.

To establish effective vaccines that produce long-term immunity, it is important to understand the genome of the SARS-CoV-2 and its functionalities. The genome of an organism is the accumulation of its entire genetic potential, encoded as a four-nucleotide sequence (Adenine-A, Guanine-G, Cytosine-C, and Thymine-T) that makes up the nucleic acids.

The COVID-19 genome is made up of single-stranded DNA.

Genome decoding is the process of determining the nucleotide sequence in a genome. Several groups have sequenced the SARS-CoV-2 genome worldwide, revealing several strains of the virus and revealing that its genome is 79 percent identical to SARS-CoV-1 and 50 percent similar to MERS-CoV.

Many facets of disease behavior remain unclear, making actionable observations difficult to come by. The use of artificial intelligence methods including sequential pattern mining (SPM), has the ability to accelerate the process of finding actionable insights and eventually contribute to a better global response.

. SPM [a form of hierarchical data mining], has been used in genomics to find patterns of unique elements in genes, study gene expression, and mine the most contiguous frequent patterns. Second, see whether the next nucleotide bases can be predicted in COVID-19 genome sequences. Third, we suggest an algorithm to find mutations in genome sequences.