Table 1: Overview of the most important properties of the different splice prediction tools. Machine learning tools are marked in light grey and deep learning tools in dark grey.

Tool	Approach	Score range	Characteristic	Training data	Input data	Interface	Source code	Year
CADD	Support vector machine with linear kernel	-	Integrates more than 60 genomic features into a single score	13,141,299 SNVs, 627,071 insertions and 926,968 deletions from simulated and observed variants	VCF file	Website, python script	Only testing	2014
DSSP	CNN with long short-term memory	0 – 1	Individual prediction for SDS and SAS	HS3D	140 nt sequence with consensus sequence the middle	Python script	Only testing	2018
GeneSplicer	Decision tree + Markov model	0 – 15	Markov model captures additional dependencies among neighboring bases at splice sites	1323 plant genes and 1115 human genes	FASTA sequence	Alamut	yes	2001
MaxEntScan	Maximum entropy	0 – 12	Use of different constraints sorted by effect on entropy, only second-order dependencies	1,821 non-redundant transcripts with 12,715 introns	9-mer FASTA sequence	Alamut	yes	2004
MMSplice	Deep learning, individual modules scoring exon, intron and splice sites	0 – 1	Predicts quantitative physical measures of splicing	Vex-seq + GENCODE	VCF file	Python package	yes	2018
MTSplice	CNN, combines MMSplice and TSplice	0 – 1	Extension of MMSplice to include tissue specific splicing	ASCOT + Vex-seq + GENCODE	VCF file	Python package	yes	2020
NNSPLICE	Hidden Markov model + neural network	0 - 1	Captures pairwise correlations between adjacent nucleotides	285 multiple-exon human DNA sequences from GenBank	FASTA sequence	Alamut	no	1997
S-CAP	Gradient-boosting tree	0 – 1	Six different models, one per region	HGMD + ClinVar + gnomAD	Variant position	txt file with Precomputed values	yes	2019

SPIDEX	Bayesian modelling	0 – 1	Tissue specific PSI values	Illumina Human Body Map 2.0 project	VCF file	Txt file with precomputed values	no	2015
SpliceAI	Deep learning with ResNet blocks	0 – 1	Considers 10,000 nucleotides, predicts nucleosome positioning from sequence	GENCODE	VCF file	Python package	yes	2019
SpliceRover	CNN	0 - 1	Identifies regions/structures of interest by normalizing contribution scores, individual models for SDS and SAS	human and plant	FASTA sequence	Website	no	2018
SpliceSiteFinder-like	Position weight matrices	0 – 100	-	-	-	Alamut	no	1987

CNN - convolutional neural network, SAS - splice acceptor site, SDS - splice donor site