Validate GTF Documentation (844659 – Simone Mallei):

The violations detected by the validator are memorized in a dictionary of errors, in a specific field depending on the type of error occurred.

In order to validate a .gtf file, you have to call the following method:
validate_file(gtf_dir, gtf_file_name) : gtf_dir will be a string that contains the path to the directory containing the file you want to validate and gtf_file_name will be the name of the file (including the '.gtf'). This method will print in the standard output the violations found by the validator and it will write them in the file named 'report-<name_file>.txt' (this time not including the '.gtf') in the directory called 'reports'.

The format constraints refer to the following documentation of the GTF Format: https://mblab.wustl.edu/GTF22.html

Error Field:	Occurrs in:	Description of what could cause the error:	File:
num_fields	Row	The number of fields in the record is not the required one (the fields should be 9).	example000.gtf
start	Row	The start value is less than 1.	example001.gtf
end	Row	The end value is less than 1.	example002.gtf
start_end	Row	The start value is greater than end value.	example003.gtf
start_codon_len	Row	The length of the start_codon record is greater than 3.	example004.gtf
stop_codon_len	Row	The length of the stop_codon record is greater than 3.	example005.gtf
feature_name	Row	The feature type in the record does not exist in the format.	example006.gtf
inter_transcript	Row	The feature ("inter" or "inter_CNS") has a non-empty transcript_id.	example007.gtf
score	Row	Neither the score value is an integer, nor a float, nor '.'.	example008.gtf
strand	Row	The strand has a value that does not exist (different from '+' and '-').	example009.gtf
frame	Row	The frame has a value not accepted (frame should be >= 0 and <= 2 in 'CDS', 'start_codon' or 'stop_codon', '.' otherwise).	example010.gtf
attributes	Row	Some attributes are not represented correctly or gene_id and transcript_id are not respectively the first two attributes.	example011.gtf
strand_interrow	Gene	The gene's records have different strands (should be only one).	example012.gtf
missing_CDS	Transcript	Even if there is at least a record about other features among 'start_codon' and 'stop_codon', there is no 'CDS' record.	example013.gtf
missing_start_codon	Transcript	Even if there is at least a record about other features among 'CDS' and 'stop_codon', there is no 'start_codon' record.	example004.gtf
emissing_stop_codon	Transcript	Even if there is at least a record about other features among 'CDS' and 'start_codon', there is no 'stop_codon' record.	example005.gtf
CDS_interrow	Transcript	At least one of the frames in the records of the transcript that has been checked is not correctly indicated or there are some intervals with some positions in common, or the length of the total CDS is not a multiple of 3.	example010.gtf
start_codon_interrow	Transcript	At least one of the frames in the records of the transcript that has been checked is not correctly indicated or there are some intervals with some positions in common, or the length of the codon is not 3.	example014.gtf
stop_codon_interrow	Transcript	At least one of the frames in the records of the transcript that has been checked is not correctly indicated or there are some intervals with some positions in common, or the length of the codon is not 3.	example015.gtf
CDS_start_codon	Transcript	The first codon of the 'CDS' is not the 'start_codon' recorded.	example016.gtf
CDS_stop_codon	Transcript	The stop codon is not positioned after the 3' base of the transcript's 'CDS'.	example017.gtf

Extra files with no violations: example018.gtf, example019.gtf

When a row violates one of the "Row" constraints, it will be ignored in the phase where the validator will search for the "Gene" and "Transcript" errors, possibly causing other violations (for example, 'missing_CDS', 'missing_start_codon', 'missing_stop_codon', or 'CDS_interrow' as well).

Once you execute the validator giving the path of a .gtf file to examine, it will write a report in a specific directory that contains only reports produced by the validator itself.

Since '5UTR' and '3UTR' are optional features, their possible violations would be ignored.