Approximate Pattern Matching Problem

**Input**: Strings *Pattern* and *Text* along with an integer *d*.

**Output**: All starting positions where *Pattern* appears as a substring of *Text* with at most *d* mismatches.
SAMPLE DATASET:

Input:
ATTCTGGA
CGCCCGAATCCAGACGCAATTCCCATATTTCCGGGACCACCTGGCGCTCCACACGGTACGGA
CGTCAATCAAAT
3

Output:
6 7 26 27

The sample dataset is not actually run on your code.
TEST DATASET 1:

Input:
AAA
TTTTTTAAATTTAAATTTTTT
2

Output:
4 5 6 7 8 11 12 13 14 15

This dataset checks if you are only counting instances where the number of mismatches is exactly equal to $d$ (i.e. ignoring instances where $mismatch < d$).
TEST DATASET 2:

Input:
GAGCGCTGG

GAGCGCTGGTTAACTCGCTACTTCCCCGACGAGCGCTGTGGCGCAAATTGGCGATGA
AACTGCAGAGAGAACTGGTCATCCAACTGAATTCTCCCCGCTATCGCATTTTGATGC
GCGCCCGCGTCGATT

2

Output:
0 30 66

This dataset checks if your code has an off-by-one error at the beginning of Text (i.e. your code is not checking the the leftmost substring of Text).
TEST DATASET 3:

Input:
AATCCTTTCA
CCAAATCCCTCATGGCATGCATTCCCGCAGTATTTTAATCCTTTCTATTCTGCATATAA
GTAGTGAAAGGTATAGAAACCGTTCAAGCCCGCAGCGGTAAAACCGGAACCATGA
TGAATGCACGCGATTCGCGGCTAATCCAACA

Output:
3 3 6 74 137

This dataset checks if your code has an off-by-one error at the end of Text (i.e. your code is not checking the the rightmost substring of Text).
TEST DATASET 4:
Input:
CCGTCATCC

CCGTCATCCGTCATCCTCGCCACGTTGGAATGCATTCCGTCATCCGTCAGGCATACT
TCTGCATATAAGTACAAACA
TGCATGTCAAAGGGAGCCCGCAGCGGTAAAACC
GAGAACCATGATGAATGCACGGCGATTGC

3

Output:
0 7 36 44 48 72 79 112

This dataset checks if your code is correctly accounting for overlapping instances of Pattern in Text.
TEST DATASET 5:

Input:
TTT
AAAAAA
3

Output:
0 1 2 3

This dataset checks if you are only counting instances of Pattern with less than d mismatches (as opposed to instances of Pattern with less than or equal to d mismatches).
TEST DATASET 6:

Input:
CCA
CCACCT
0

Output:
0

This dataset checks if your code works with input where $d = 0$ (i.e. only perfect matches are allowed).